COUNTWAY LIBRARY
HC 1GEA R

A STUDY OF AMERICAN INTELLIGENCE

BY

CARL C. BRIGHAM

Digitized by the Internet Archive
in 2026

https://archive.org/details/studyofamericani0000carl

A STUDY OF AMERICAN INTELLIGENCE

A STUDY OF AMERICAN INTELLIGENCE

BY

CARL C. BRIGHAM, PH.D.
ASSISTANT PROFESSOR OF PSYCHOLOGY
IN PRINCETON UNIVERSITY

A FOREWORD
BY ROBERT M. YERKES, PH.D.
CHAIRMAN RESEARCH INFORMATION
SERVICE: NATIONAL RESEARCH COUNCIL

PRINCETON
PRINCETON UNIVERSITY PRESS
LONDON: HUMPHREY MILFORD
OXFORD UNIVERSITY PRESS
1923

BOSTON MEDICAL LIBRARY
IN THE
FRANCIS A. COUNTWAY
LIBRARY OF MEDICINE

COPYRIGHT 1922 BY CARL C. BRIGHAM

PRINTED AT THE PRINCETON UNIVERSITY PRESS, PRINCETON, N. J.

FOREWORD

Two extraordinarily important tasks confront our nation: the protection and improvement of the moral, mental and physical quality of its people and the re-shaping of its industrial system so that it shall promote justice and encourage creative and productive workmanship. I have been asked to write this Foreword because of my official connection, as chief of the Division of Psychology, Office of the Surgeon General of the Army, with psychological examining during the war, but I have consented to write it because of my intense interest in the practical problems of immigration and my conviction that the psychological data obtained in the army have important bearing on some of them.

When in April, 1917, I visited Canada to learn what use our neighbors were making of psychological principles and methods in their military activities, I found Mr. Carl C. Brigham attached as psychologist to the Military Hospitals Commission. With him as my guide, I spent several hours in interviewing military and civil officers and in discussing our mutual problems and needs. The valuable information which Mr. Brigham helped me to secure and his advice contributed substantially to the report which I later presented to my professional colleagues at home, and to representatives of the United States army.

In October, 1917, our friend, eager for larger opportunities for professional service than the Canadian army promised, accepted appointment in the Sanitary Corps of the United States army for psychological service. He aided

efficiently in the trials of methods of examining at Camp Dix, New Jersey, and he was then ordered to the office of the Surgeon General in Washington to help with the revision of tests and the preparation of new methods. Thus he became thoroughly familiar with the procedures and results of psychological examining in the army, while at the same time contributing generously of ideas, labor and enthusiasm. With deep satisfaction I use this opportunity to mention Mr. Brigham's national service and his exceptional fitness to study and to discuss the relations of army measurements of intelligence to nativity and residence.

It appears that Mr. Charles W. Gould, a clear, vigorous, fearless thinker on problems of race characteristics, amalgamation of peoples and immigration, raised perplexing questions which drove Mr. Brigham to his careful and critical re-examination, analysis, and discussion of army data concerning the relations of intelligence to nativity and length of residence in the United States. In a recently published book, *America, A Family Matter*, to which this little book is a companion volume, Mr. Gould has pointed the lessons of history for our nation and has argued strongly for pure-bred races.

For the observational data which Mr. Brigham used in preparing this book we are indebted to the competent and devoted company of psychologists which during the war labored in camp and laboratory on the preparation of methods, the conduct of examinations, and the application of results. But the fruits of the labors of these many psychologists might have been lost to the world had it not been for the insight, zeal, and industry of Carl R. Brown, Mark A. May and Edwin G. Boring, who evolved methods of statistical treatment, applied them and prepared the resulting materials for publication.

Mr. Brigham has rendered a notable service to psychol-

ogy, to sociology, and above all to our law-makers by carefully re-examining and re-presenting with illuminating discussion the data relative to intelligence and nativity first published in the official report of psychological examining in the United States army. Far from belittling or casting doubt on the general reliability of the results contained in the report, he has essentially confirmed the major findings in the field of his special inquiry and has adduced new evidences of the trustworthiness and scientific value of the statistical methods used by military psychologists. His task has been arduous and difficult, involving an immense amount of tedious labor for mathematical calculations and critical study of results. The volume which is the outcome of Mr. Brigham's inquiry, and which I now have the responsibility and satisfaction of recommending, is substantial as to fact and important in its practical implications. It is not light or easy reading but it is better worth re-reading and reflective pondering than any explicit discussion of immigration which I happen to know. The author presents not theories or opinions but facts. It behooves us to consider their reliability and their meaning, for no one of us as a citizen can afford to ignore the menace of race deterioration or the evident relations of immigration to national progress and welfare.

Robert M. Yerkes.

Washington, D. C.
June 1922

CONTENTS

		PAGE
FOREWORD		v
INTRODUCTION		xix
PART I : THE ARMY TESTS		
Section 1	EXAMINATION ALPHA	3
Section 2	EXAMINATION BETA	32
Section 3	THE INDIVIDUAL EXAMINATIONS	54
Section 4	RELIABILITY OF THE MEASURES	59
PART II : STATISTICAL ANALYSIS OF THE ARMY TEST RESULTS		
Section 1	THE PRINCIPAL SAMPLE	75
Section 2	ANALYSIS OF THE MAIN GROUPS OF THE PRINCIPAL SAMPLE	77
Section 3	ANALYSIS OF THE WHITE DRAFT INTO FOREIGN AND NATIVE BORN	84
Section 4	ANALYSIS OF THE FOREIGN BORN WHITE DRAFT INTO YEARS OF RESIDENCE GROUPS	88
Section 5	ANALYSIS OF IMMIGRATION TO THE UNITED STATES	112
Section 6	ANALYSIS OF THE FOREIGN BORN WHITE DRAFT BY COUNTRY OF BIRTH	118
Section 7	RELIABILITY OF THE RESULTS	154
Section 8	THE RACE HYPOTHESIS	157
Section 9	RE-EXAMINATION OF PREVIOUS CONCLUSIONS IN THE LIGHT OF THE RACE HYPOTHESIS	177
Section 10	COMPARISON OF OUR RESULTS WITH THE CONCLUSIONS OF OTHER WRITERS ON THE SUBJECT	182
CONCLUSIONS		197

PLATES

PLATE		
I	*Alpha test* 1	: ORAL DIRECTIONS
II	*Alpha test* 2	: ARITHMETICAL REASONING
III	*Alpha test* 3	: PRACTICAL JUDGMENT
IV	*Alpha test* 4	: SYNONYM-ANTONYM
V	*Alpha test* 5	: DISARRANGED SENTENCES
VI	*Alpha test* 6	: NUMBER SERIES COMPLETION
VII	*Alpha test* 7	: ANALOGIES
VIII	*Alpha test* 8	: INFORMATION
IX	*Beta test* 1	: MAZE
X	*Beta test* 2	: CUBE ANALYSIS
XI	*Beta test* 3	: X-O SERIES
XII	*Beta test* 4	: DIGIT-SYMBOL
XIII	*Beta test* 5	: NUMBER CHECKING
XIV	*Beta test* 6	: PICTURE COMPLETION
XV	*Beta test* 7	: GEOMETRICAL CONSTRUCTION

FIGURES

FIGURE		PAGE
1.	Distribution of scores of the oral directions test	6
2.	The Gaussian normal distribution	7
3.	Distribution of scores of the arithmetical reasoning test	10
4.	Distribution of scores of the practical judgment test	14
5.	Distribution of scores of the synonym-antonym test	17
6.	Distribution of scores of the disarranged sentence test	20
7.	Distribution of scores of the number series completion test	23
8.	Distribution of scores of the analogies test	25
9.	Distribution of scores of the information test	28
10.	Black-board chart for demonstrating the maze test	33
11.	Distribution of scores of the maze test	34
12.	Black-board chart for demonstrating the cube analysis test	36
13.	Distribution of scores of the cube analysis test	37
14.	Black-board chart for demonstrating the X-O series test	39
15.	Distribution of scores of the X-O series test	40
16.	Black-board chart for demonstrating the digit-symbol test	42

FIGURE PAGE

17. Distribution of scores of the digit-symbol test 43
18. Black-board chart for demonstrating the number checking test 45
19. Distribution of scores of the number checking test 46
20. Black-board chart for demonstrating the picture completion test 48
21. Distribution of scores of the picture completion test 49
22. Black-board chart for demonstrating the geometrical construction test 51
23. Distribution of scores of the geometrical construction test 52
24. The normal distribution curve 59
25. A skewed distribution curve 60
26. Examination alpha as independent of education 65
27. Distribution of intelligence scores according to rank 66
28. Success in Officers' Training Camps as predicted by examination alpha 67
29. Comparison of army test records with various independent criteria 69
30. Success in civil occupations compared with army test records 70
31. Distributions of scores of the white officers, white draft, and negro draft on the combined scale 81
32. Distributions of scores of the native born and foreign born white draft on the combined scale 87
33. Apparently increasing average intelligence with increasing years of residence 94
34. Distributions of the alpha scores of three groups 108
35. Analysis of immigration by countries 114
36. Relative standing of the nativity groups according to their average intelligence 124

FIGURE PAGE

37. Relative standing of the nativity groups in the proportions of A and B men, and D, D− and E men 146
38. The proportion of each nativity group obtaining scores at or above the average of the white officers 149
39. The proportion of each nativity group at or below the average of the negro draft 151
40. The proportion of each nativity group testing below the approximate "mental age" of eight 153
41. Analysis of immigration to the United States according to the estimated amount of Nordic, Mediterranean and Alpine blood 164
42. Volume of immigration by decades 166
43. The distributions of the intelligence scores of the Nordic, Mediterranean and Alpine groups 170
44. The distributions of the intelligence scores of the English speaking Nordic and the non-English speaking Nordic groups 173
45. The distributions of the intelligence scores of the non-English speaking Nordic group and the combined Mediterranean and Alpine groups 175
46. The decline of intelligence with each succeeding period of immigration 198
47. The constituent elements of American intelligence 200

TABLES

		PAGE
1.	Distribution of the intelligence scores of the main groups of the principal sample on the combined scale	80
2.	Analysis of the white draft into foreign born and native born groups	86
3.	Analysis of the foreign born white draft by years of residence in the United States	90
4.	Comparison of the average scores on the combined scale of the five years of residence groups of the foreign born white draft	91
5.	Comparison of the average scores on the combined scale of the native born white draft with the five years of residence groups of the foreign born white draft	92
6.	Per cent. that emigration was of immigration for 15 countries since 1908	98-99
7.	Distribution of alpha scores of five groups	106
8.	Per cent. of total immigration coming from various countries during periods roughly corresponding to the five years of residence groups	113
9.	Analysis of the foreign born white draft by country of birth: actual distributions	120-121
10.	Analysis of the foreign born white draft by country of birth: percentage distributions	122-123
11.	Differences between England and other countries	126
12.	Differences between Scotland and other countries	127
13.	Differences between Holland and other countries	128
14.	Differences between Germany and other countries	129
15.	Differences between the United States and other countries	130
16.	Differences between Denmark and other countries	131
17.	Differences between Canada and other countries	132

PAGE
18. Differences between Sweden and other countries 133
19. Differences between Norway and other countries 134
20. Differences between Belgium and other countries 135
21. Differences between Ireland and other countries 136
22. Differences between Austria and other countries 137
23. Differences between Turkey and other countries 138
24. Differences between Greece and other countries 139
25. Differences between Russia and other countries 140
26. Differences between Italy and other countries 141
27. Differences between Poland and other countries 142
28. Per cent. of each nativity group in the A and B groups 144
29. Per cent. of each nativity group in the D, D− and E groups 145
30. Per cent. of each nativity group at or above the average of the white officers 148
31. Per cent. of each nativity group at or below the average of the negro draft 150
32. Per cent. of each nativity group below the approximate "mental age" of eight 152
33. Tentative estimates of the proportion of Nordic, Alpine and Mediterranean blood in each of the European countries 159
34. Arrivals of alien passengers and immigrants, 1820 to 1920 160-161
35. Estimate of the amount of Nordic, Mediterranean and Alpine blood coming to this country from Europe in each decade since 1840 163
36. Analysis of the foreign born white draft by races 169
37. Analysis of the Nordic sample into an English speaking Nordic group and a non-English speaking Nordic group 172
38. Population of the United States in 1920 203

ACKNOWLEDGMENTS

This study is a continuation of the work of the small group of psychologists who carried on the difficult task of analyzing the data from the army psychological examinations in the office of the Surgeon General of the Army. My presentation contains nothing new in methodology and is merely an extension of lines of investigation suggested by this group of workers. It rests on the foundations which they built.

I wish to make especial acknowledgment to Colonel Robert M. Yerkes, who has read the manuscript several times in its various stages of preparation and has given many helpful suggestions. Professor Carl R. Brown of the University of Michigan, formerly of the Surgeon General's staff, assisted me when I first began to use the combined scale, and subsequently read Sections 1 to 7 of Part II in manuscript. Professor E. G. Boring of Harvard University read Sections 1 to 7 of Part II, and gave me invaluable assistance, especially in the treatment of Section 4. Professor Mark A. May of Syracuse University, Professors Edwin G. Conklin and Howard C. Warren of Princeton University read Sections 1 to 7 of Part II, and suggested many important changes. Without the assistance of all of these gentlemen I could not have carried through the task.

Mr. Charles W. Gould suggested this continuation of the army investigations in the first instance, has sponsored the work throughout, has read and re-read all of the manuscript at every stage of its preparation, and is mainly responsible for the whole work. In my treatment of the race hypothesis I have relied on his judgment and on two books, Mr. Madison Grant's *Passing of the Great Race*, and

Professor William Z. Ripley's *Races of Europe*. These three gentlemen cannot be held responsible however for my percentage analysis of the present racial constitution of European countries, an analysis which I made as a novice in the field of anthropology, and for which I offer further apologies in the text of Section 8.

Mr. David M. Maynard and Mr. Charles H. Helliwell, two of my undergraduate students, assisted me faithfully in carrying through the laborious statistical calculations involved in using the combined scale.

CARL C. BRIGHAM

Princeton, N. J.
September 1922

INTRODUCTION

The question of the differences that may exist between the various races of man, or between various sub-species of the same race, or between political aggregations of men in nationality groups may easily become the subject of the most acrimonious discussion. The anthropologists of France and Germany, shortly after the close of the Franco-Prussian war, fought another national war on a small scale. It is difficult to keep racial hatreds and antipathies out of the most scholarly investigations in this field. The debate becomes especially bitter when mental traits are discussed. No one can become very indignant on finding his race classified by its skull dimensions, stature, or hair color, but let a person discover the statement that his race is unintelligent or emotionally unstable, and he is immediately ready to do battle.

Until recent years we have had no methods available for measuring mental traits scientifically, so that the literature on race differences consists largely of opinions of students who are very apt to become biased, when, leaving the solid realm of physical measurements, they enter the more intangible field of estimating mental capacity.

Gradually, however, various investigators using more or less refined psychological measurements commenced to assemble a body of data that will some day reach respectable proportions. The status of the psychological investigations of race differences up to 1910 has been admirably summarized by Woodworth.[1] Since 1910, we have witnessed

[1]R. S. Woodworth. *Racial Differences in Mental Traits*, Science, New Series, Vol. 31, pp. 171–186.

in this country a remarkable development in methods of intelligence testing, and these methods have been applied to the study of race differences. Scattered investigations report and compare the intelligence scores of children of white, negro, or Indian parentage, and sometimes the scores of various nationality or nativity groups. The results of these investigations are, however, almost impossible to correlate, for they have been made by different methods, by different measuring scales, on children of a wide variety of chronological ages, and above all, on comparatively small groups of subjects, so that conclusions based on the studies have no high degree of reliability.

For our purposes in this country, the army mental tests give us an opportunity for a national inventory of our own mental capacity, and the mental capacity of those we have invited to live with us. We find reported in Memoir XV of the National Academy of Sciences[1] the intelligence scores of about 81,000 native born Americans, 12,000 foreign born individuals, and 23,000 negroes. From the standpoint of the numbers examined, we have here an investigation which, of course, surpasses in reliability all preceding investigations, assembled and correlated, a hundred fold. These army data constitute the first really significant contribution to the study of race differences in mental traits. They give us a scientific basis for our conclusions.

When we consider the history of man during the half million years which have probably elapsed since the time of the erect primate, *Pithecanthropus*, the temporary political organizations, such as Greece, Rome, and our modern national groups, become of minor importance compared with the movements of races and peoples that have occurred. The tremendous expansion of the Alpine race at the end of the Neolithic and the beginning of the Bronze

[1] *Psychological Examining in the United States Army.* Edited by Robert M. Yerkes. Washington: Government Printing Office, 1921, pp. 890.

Period, the submergence of this race by the Nordics in the 2000 years preceding the Christian era, and the subsequent peaceful re-conquest of Eastern Europe by the Alpine Slavs from the Dark Ages on, represent an historical movement in comparison with which the Great World War of 1914 resembles a petty family squabble.

If the history of the United States could be written in terms of the movements of European peoples to this continent, the first stage represents a Nordic immigration, for New England in Colonial times was populated by an almost pure Nordic type. There followed then a period of Nordic expansion. The next great movement consisted of the migrations of Western European Mediterraneans and Alpines from Ireland and Germany, a movement which started about 1840, and which had practically stopped by 1890. Since there is a considerable proportion of Nordic blood in Ireland and Germany, we should not regard the original Nordic immigration as a movement which stopped suddenly, but merely as having dwindled to two-fifths or one-half of the total racial stock coming here between 1840 and 1890. The third and last great movement consisted of migrations of the Alpine Slav and the Southern European Mediterraneans to this continent, a movement that started about 1890, and which has not yet ceased. Running parallel with the movements of these European peoples, we have the most sinister development in the history of this continent, the importation of the negro.

The army mental tests enable us to analyze the elements entering into American intelligence. The intelligence test records of the native born, the foreign born, and the negro are at our disposal. The records deserve the most serious study. But before considering the results of the army tests, a person should be well informed concerning the nature of the tests, and the manner in which they were constructed.

The army psychological tests included three types of examination:

(1) Group examination alpha, which included eight different sorts of tests, most of which involved the ability to read English.

(2) Group examination beta, which included seven different sorts of tests, none of which involved the ability either to read English or to understand spoken English, the tests consisting of pictures, designs, etc., and being given by instructions in pantomime.

(3) Individual examinations of two types:
   (*a*) Those involving the use of English, the Stanford revision of the Binet-Simon scale and the point scale, and
   (*b*) Those involving no English, consisting of construction puzzles, etc., the instructions being given by gestures,—the "performance scale."

When a detachment reported for psychological examination, the first step was that of separating the English speaking and literate from the non-English speaking or illiterate. Those who were both English speaking and literate were given examination alpha. All others were sent to beta. At the close of examination alpha, all men who had made low scores were sent to beta. After examination beta had been given, the examiners tried to recall for individual examinations all men who had made a low score in beta. In the rush of examining it was impossible to recall all men for individual examinations who should have been given special examinations, and some men were graded on alpha who should have been graded on beta, and vice versa, but

most men were properly graded by the rough methods in use. In each one of the examinations the range of scores was so great that most men had an opportunity to score.

The great contribution of the committee that first devised the army examining methods and of the men who subsequently developed additional methods in the army consisted of creating and standardizing group examinations alpha and beta. The methods of individual examining were already in existence, the Stanford-Binet scale being an elaboration of Binet's "mental age" scale, and the tests of the performance scale having been more or less completely worked out by other investigators. The task of examining men in large groups was first carried through successfully in the army. Before the war, many psychologists would have scoffed at the notion of examining two or three hundred men at once by giving them booklets containing different sorts of tests, but the large group examinations became matters of daily routine. Group tests have subsequently been tried out in schools and industries with excellent results from the standpoint of test administration. Indeed, when the army alpha examination was given at Ohio State University in October, 1919, practically the entire student body, 6000 in number, was tested by five examiners in eight hours. In the service, it was found that one examiner could control a group of 200 men with ease. The alpha instructions were read by the examiner, and the men ordered to start and stop at the proper time. Examination beta was more difficult to administer, and was given to smaller groups.

The statistical methods of treating the results of the army tests used in this study are rather intricate, but the principles involved are easily understood. At the outset we must frankly admit that there were minor errors in the three types of examinations given. We can not correct the

type of tests that were used, but we can correct the method of scoring them. Most of the difficulties of scoring arise from the fact that different types of measuring scales were used. During the war, the different scales were converted into one general scale of letter grades (A, B, C+, C, C−, D and D−). This method was rough, and although it answered the purposes of the army at the time, it can not be used in any scientific interpretation of the results.

Examination alpha was scored by finding the score on each of the eight tests, adding to get a total, and then converting the total into a letter grade. Beta was similarly scored. It is apparent that some tests in alpha might be more difficult than others, that some tests in beta might be easier than any test in alpha, and that variations might have occurred which it was impossible to predict at the time the examinations were made. Recognizing these facts, then, the army statisticians worked out another method of scoring the results, which eliminates all of these sources of error. This method is known as the *combined scale*, a theoretical intelligence scale running from 0 to 25, into which the alpha, beta and individual examination scores may be converted, so that we finally have one measurement instead of three.

Psychological measurements involve much more than creating tests and giving tests. After all the results are in, we still have the problem of interpreting the results, and this interpretation is largely a statistical problem. Too much credit can not be given to the staff of the Psychological Division of the Surgeon General's Office, who continued in the service long after the war was over, patiently studying and analyzing the results. The combined scale was very largely the work of two young psychologists, Carl R. Brown and Mark A. May, and their work on this problem, reported in Chapter 2, Part 3 of Memoir XV, is with-

out doubt the greatest contribution that has yet been made to the statistical phases of the science of mental measurement.

The theory underlying the combined scale is simply that of regarding each test of alpha and beta as a separate measuring scale. One group of individuals including 1047 men born in English speaking countries, was examined on alpha, re-examined on beta, and if possible, examined again on the Stanford-Binet scale. This group of 1047 cases constituted the basis on which a method of combining the separate tests into a combined scale was empirically evolved.

From now on in the course of our study of the army test records, we must regard alpha and beta as two booklets containing, in all, fifteen different measuring scales of intelligence. The first step in the study consists of understanding the nature of each of the fifteen scales. In Part I, the fifteen tests have been reproduced (Plates I to XV), and the actual records of the 1047 men shown in each instance, so that the reader may see exactly how the tests worked.

PART I

THE ARMY TESTS

SECTION I

EXAMINATION ALPHA

Alpha Test 1. Oral Directions

The first test in alpha consisted of a series of commands or directions which were to be executed quickly. The instructions, with the incidental commands about stopping and starting eliminated, are reproduced below. One may read the instructions for each item to himself slowly and turn the page to Plate I to test his own ability to execute the commands.

Instructions : Oral Directions (Form 8)

Item 1. Time limit : 5 seconds.
"Make a figure 2 in the second circle and also a cross in the third circle."

Item 2. Time limit : 5 seconds.
"Draw a line from circle 1 to circle 4 that will pass *below* circle 2 and *above* circle 3."

Item 3. Time limit : 10 seconds.
"Make a figure 1 in the space which is in the square but not in the triangle, and also make a cross in the space which is in the triangle and in the square."

Item 4. Time limit : 10 seconds.
"Make a figure 2 in the space which is in the circle but not in the triangle or square, and also make a figure 3 in the space which is in the triangle and circle, but not in the square."

Item 5. Time limit : 10 seconds.
"If taps sound in the evening, then put a cross

in the first circle; if not, draw a line *under* the word NO."

ITEM 6. Time limit : 10 seconds.
"Put in the first circle the right answer to the question: 'How many months has a year?' In the second circle do nothing, but in the fifth circle put any number that is a wrong answer to the question that you have just answered correctly."

ITEM 7. Time limit : 10 seconds.
"*Cross out* the letter just after F and also draw a line *under* the second letter after I."

ITEM 8. Time limit : 10 seconds.
"Make in the *first* circle the *last* letter of the *first* word; in the *second* circle the *middle* letter of the second word and in the *third* circle the *first* letter of the *third* word."

ITEM 9. Time limit : 15 seconds.
"*Cross out* each number that is more than 50 but less than 60."

ITEM 10. Time limit : 15 seconds.
"Put a 4 or a 5 in each of the two largest parts and any number between 6 and 9 in the part next in size to the smallest part."

ITEM 11. Time limit : 25 seconds.
"Draw a line through every odd number that is not in a square, and also through every odd number that is in a square with a letter."

ITEM 12. Time limit : 10 seconds.
"If 4 is more than 2, then cross out the number 3 unless 3 is more than 5, in which case draw a line *under* the number 4."

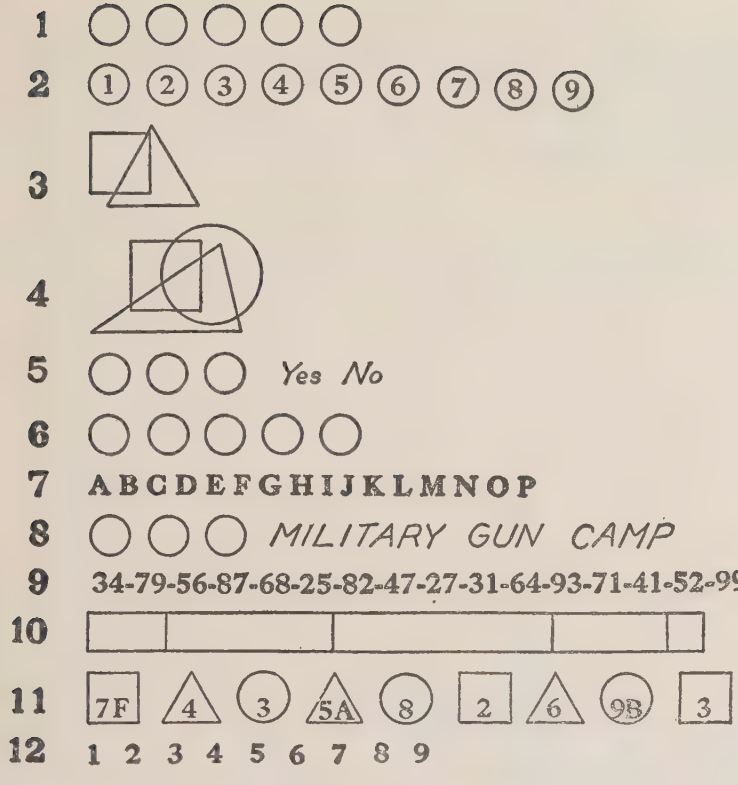

Plate I. Alpha Test 1: Oral Directions.

In scoring the papers one point was given for each correct response. The group of 1047 individuals born in English speaking countries obtained the following scores:

Total score of test No.	0	1	2	3	4	5	6	7	8	9	10	11	12
Number who made each score	73	78	93	116	100	121	131	94	82	67	52	28	13

These scores are shown graphically in Figure 1, the horizontal direction indicating the total score from the lowest possible (0) to the highest possible (12), while

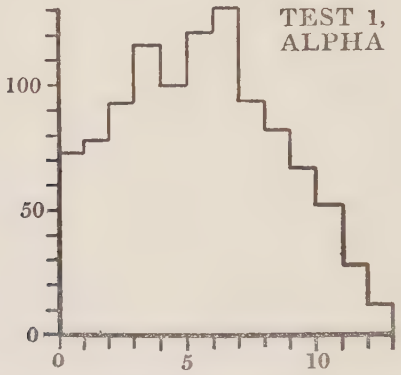

Figure 1. Distribution of scores of the Oral Directions test. (From p. 624, *Memoir XV.*)

the vertical scale represents the number of cases getting each score, 72 at 0, 78 at 1, etc.

For our purposes, we do not want a test that everyone can pass, for if everyone passed, no one would be graded. An ideal test would be one in which practically everyone could obtain some score and which very few could finish.

Then all people would be measured. An ideal test would also show a distribution of responses grouped symmetrically about the average, for, as a general rule, all measures of individual differences in mental traits show a distribution similar to the normal probability or chance distribution. An ideal test would give the type of distribution shown in Figure 2.

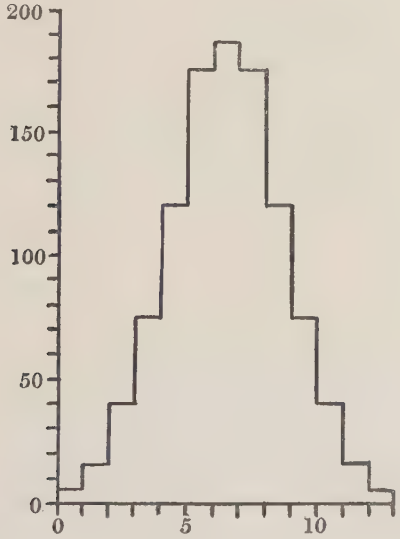

Figure 2. The Gaussian normal distribution.

Examining Figure 1, one may see that in general the oral directions test gives a type of distribution which is approximately similar to the Gaussian normal distribution shown in Figure 2. Our distribution is limited slightly at the lower end, and it is easy to imagine that the introduction of two

or three items easier than any in the present test would give us a step-down at the zero end of the scale similar to the one at the upper end.

The oral directions test really gives an excellent score distribution. It is not a "speed" test in the popular sense, for the time limits for each item, while short, give ample time for following the directions. If a person understands the directions he can execute them easily in the time allowed. If the directions are not understood, an hour to execute them is no more generous than five seconds. In practice this test was useful in "acclimating" the recruit to the conditions of the examination. It was probably one of the poorest tests in alpha as a genuine test of intelligence, but it served its purpose as a "warming up" test. It is an adaptation of a type of test that has been used in psychological laboratories for many years with rather mediocre results.

Alpha Test 2.—Arithmetical Problems

Time limit : 5 minutes

Test 2 is more of a reasoning test than a measure of proficiency in the fundamental arithmetical operations. The first items really constitute a literacy test, for if a person can read, he can answer the questions correctly. The distribution of scores in this test is shown in Figure 3.

The zero scores (66 in number) are probably due to the inclusion of illiterates in the group of 1047 cases. Disregarding the zero scores, the distribution is regular. This test illustrates admirably the principle of fixing a time limit such that very few people can answer all the items correctly. The approximate rule adopted in fixing the time limit in the first instance was that this limit should be such that not more than five per cent. of an unselected group

Get the answers to these examples as quickly as you can.
Use the side of this page to figure on if you need to.

SAMPLES
1 How many are 5 men and 10 men?.....Answer (15)
2 If you walk 4 miles an hour for 3 hours, how far do you walk?.....Answer (12)

1 How many are 60 guns and 5 guns?.....Answer ( )
2 If you save $9 a month for 3 months, how much will you save?.....Answer ( )
3 If 48 men are divided into squads of 8, how many squads will there be?.....Answer ( )
4 Mike had 11 cigars. He bought 2 more and then smoked 7. How many cigars did he have left?.....Answer ( )
5 A company advanced 8 miles and retreated 2 miles. How far was it then from its first position?.....Answer ( )
6 How many hours will it take a truck to go 42 miles at the rate of 3 miles an hour?.....Answer ( )
7 How many pencils can you buy for 60 cents at the rate of 2 for 5 cents?.....Answer ( )
8 A regiment marched 40 miles in five days. The first day they marched 9 miles, the second day 6 miles, the third 10 miles, the fourth 6 miles. How many miles did they march the last day?.....Answer ( )
9 If you buy 2 packages of tobacco at 8 cents each and a pipe for 65 cents, how much change should you get from a two-dollar bill?.....Answer ( )
10 If it takes 4 men 3 days to dig a 120-foot drain, how many men are needed to dig it in half a day?.....Answer ( )
11 A dealer bought some mules for $2,000. He sold them for $2,400, making $50 on each mule. How many mules were there?.....Answer ( )
12 A rectangular bin holds 200 cubic feet of lime. If the bin is 10 feet long and 5 feet wide, how deep is it?.....Answer ( )
13 A recruit spent one-eighth of his spare change for post cards and twice as much for a box of letter paper, and then had $1.00 left. How much money did he have at first.....Answer ( )
14 If 3½ tons of clover cost $14, what will 6½ tons cost?..Answer ( )
15 A ship has provisions to last her crew of 700 men 2 months. How long would it last 400 men?.....Answer ( )
16 If an aeroplane goes 250 yards in 10 seconds, how many feet does it go in a fifth of a second?.....Answer ( )
17 A U-boat makes 8 miles an hour under water and 20 miles on the surface. How long will it take to cross a 100-mile channel, if it has to go two-fifths of the way under water?.....Answer ( )
18 If 134 squads of men are to dig 3,618 yards of trench, how many yards must be dug by each squad?.....Answer ( )
19 A certain division contains 5,000 artillery, 15,000 infantry, and 1,000 cavalry. If each branch is expanded proportionately until there are in all 23,100 men, how many will be added to the artillery?.....Answer ( )
20 A commission house which had already supplied 1,897 barrels of apples to a cantonment delivered the remainder of its stock to 37 mess halls. Of this remainder each mess hall received 54 barrels. What was the total number of barrels supplied?..Answer ( )

Plate II. Alpha Test 2: Arithmetical Reasoning (Form 8).

would complete all the items in a test. In our group of 1047 cases, 5 persons answered 18 problems correctly in the five minutes allowed, but no one answered more than 18 problems correctly. Of course no one was expected to answer them all. If a person passed 65% of each test in alpha he was graded "A"; perfection was not required.

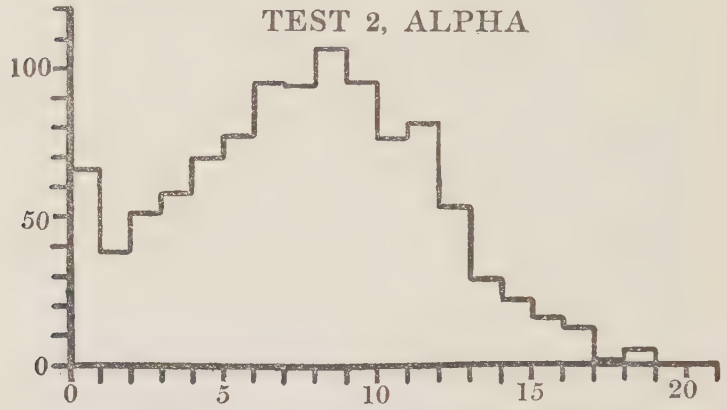

Figure 3. Distribution of scores of the Arithmetical Reasoning test. (From p. 624, *Memoir XV.*)

One often hears the statement that the army tests were "speed" tests, and penalized the slow but accurate individual. Experiments were made to determine how the results would change with extended time. A group of 475 men examined showed in Test 2 an improvement from an average of 8.00 to 9.16 with double time. In five minutes they solved on an average 8 problems correctly, in ten minutes 9.16. The relationship between single time and double time scores may be measured by the statistical value known as the coefficient of correlation. Two measures that stand in a perfect one to one correspondence have a coefficient

of correlation of 1.00. Two measures that stand in a perfect chance relationship have a correlation coefficient of 0. In practice it is found that a correlation of 0.90 is so high that one might substitute one series of measures for the other without seriously changing the results. The correlation between the single time and double time scores was 0.937, a value so high that it indicates that there were very few changes in the relative position of the members of the group, and that such changes as occurred were small.

The experiments that were conducted on time limits with the various tests all pointed to the conclusion that the results would not be changed with the more extended time limits. Of course the absolute scores would be higher with the extended time, but the relative position of the members of the group would be about the same. In the experiment on double time referred to above, all the tests from 2 to 8 in alpha showed coefficients of correlation between single time and double time above 0.90 except Test 3 (0.879), and the correlation of the two total scores obtained under single and double time was 0.967. The army experimenters after considering all the evidence concluded that "doubling the time does not result in any demonstratable improvement in alpha as a whole." (p. 417). It is probably true that very high scores depend on "speed," but inasmuch as a person only needed to answer correctly 65% of the items to be rated "A" and 50% of the items to be rated "B," it can not be considered that "speed" is a factor that would affect the results seriously.

The army findings of a correlation of 0.967 between the single time and double time trials of alpha, and the general conclusion that the results would not have been changed appreciably with more liberal time allowances, definitely controvert the popular belief that anything which is performed with a time limit handicaps the "slow but sure"

individual. Popular judgment classifies the population into two groups, the "slow but sure" and the "quick and inaccurate," and would have us believe that the quick type must of necessity be inaccurate, and that the sluggish individual is infallible. Science shows us that we really rate individuals on two scales, a scale of speed and a scale of accuracy, and that we find people who are both quick and accurate as well as people who are slow and inaccurate. Science would elaborate the popular classification by adding these two types. The popular "slow but sure" characterization is more apt to be an apology for dullness than a scientific diagnosis. At least in our consideration of the army test results we may definitely discard the opinion that we are testing "speed" rather than intelligence. The arithmetical reasoning test in alpha actually proved to be one of the best tests in the series.

Alpha Test 3.—Practical Judgment

Time limit : 1½ minutes

The practical judgment test is one of the most interesting tests in alpha from many standpoints. There is no other test in alpha which contains, in all of the five forms used, so many individual items that may be criticised by a person who actually inquires into the logical validity of the answers accepted as correct. Item 12, for instance, might profitably be taken as the subject of an intercollegiate debate, as it has been the subject of many debates in the history of penology. The critics of the army tests are all too apt to consider the whole scale invalid if they can discover a single incorrect item, for they fail to realize that a person could fall down on 35% of the individual items and still be rated "A."

This is a test of common sense. Below are sixteen questions. Three answers are given to each question. You are to look at the answers carefully; then make a cross in the square before the best answer to each question, as in the sample:

SAMPLE
Why do we use stoves? Because
- ☐ they look well
- ☒ they keep us warm
- ☐ they are black

Here the second answer is the best one and is marked with a cross. Begin with No. 1 and keep on until time is called.

1 It is wiser to put some money aside and not spend it all, so that you may
- ☐ prepare for old age or sickness
- ☐ collect all the different kinds of money
- ☐ gamble when you wish

2 Shoes are made of leather, because
- ☐ it is tanned
- ☐ it is tough, pliable and warm
- ☐ it can be blackened

3 Why do soldiers wear wrist watches rather than pocket watches? Because
- ☐ they keep better time
- ☐ they are harder to break
- ☐ they are handier

4 The main reason why stone is used for building purposes is because
- ☐ it makes a good appearance
- ☐ it is strong and lasting
- ☐ it is heavy

5 Why is beef better food than cabbage? Because
- ☐ it tastes better
- ☐ it is more nourishing
- ☐ it is harder to obtain

6 If some one does you a favor, what should you do?
- ☐ try to forget it
- ☐ steal for him if he asks you to
- ☐ return the favor

7 If you do not get a letter from home, which you know was written, it may be because
- ☐ it was lost in the mails
- ☐ you forgot to tell your people to write
- ☐ the postal service has been discontinued

8 The main thing the farmers do is to
- ☐ supply luxuries
- ☐ make work for the unemployed
- ☐ feed the nation

☞ Go to No. 9 above

9 If a man who can't swim should fall into a river, he should
- ☐ yell for help and try to scramble out
- ☐ dive to the bottom and crawl out
- ☐ lie on his back and float

10 Glass insulators are used to fasten telegraph wires because
- ☐ the glass keeps the pole from being burned
- ☐ the glass keeps the current from escaping
- ☐ the glass is cheap and attractive

11 If your load of coal gets stuck in the mud, what should you do?
- ☐ leave it there
- ☐ get more horses or men to pull it out
- ☐ throw off the load

12 Why are criminals locked up?
- ☐ to protect society
- ☐ to get even with them
- ☐ to make them work

13 Why should a married man have his life insured? Because
- ☐ death may come at any time
- ☐ insurance companies are usually honest
- ☐ his family will not then suffer if he dies

14 In Leap Year February has 29 days because
- ☐ February is a short month
- ☐ some people are born on February 29th
- ☐ otherwise the calendar would not come out right

15 If you are held up and robbed in a strange city, you should
- ☐ apply to the police for help
- ☐ ask the first man you meet for money to get home
- ☐ borrow some money at a bank

16 Why should we have Congressmen? Because
- ☐ the people must be ruled
- ☐ it insures truly representative government
- ☐ the people are too many to meet and make their laws

Plate III. Alpha Test 3: Practical Judgment (Form 8).

The distribution of the scores made in Test 3 is shown in Figure 4.

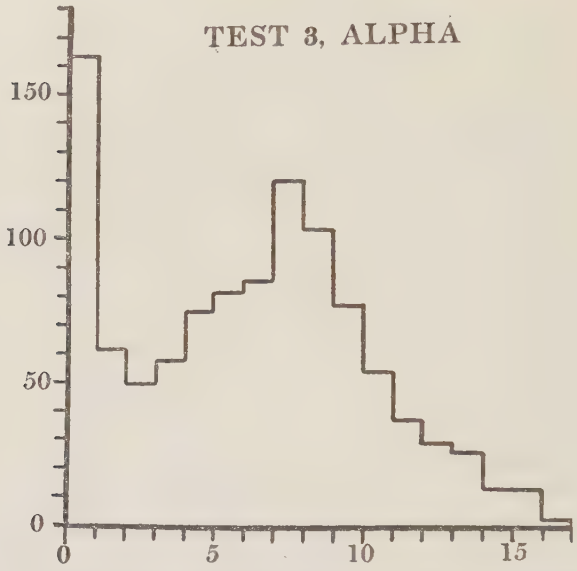

Figure 4. Distribution of scores of the Practical Judgment test. (From p. 624, *Memoir XV.*)

Disregarding the large number of zero scores (163), which are probably the result of illiteracy plus failure to understand instructions, and also recognizing the fact that a few low positive scores may be due to chance, we may look at the distribution as entirely satisfactory.

Many persons object to the short time limit (1½ minutes), but the test was undoubtedly more effective with this short limit than it would have been with the longer time.

The average score improves from 6.32 to 9.85 with double time, and the correlation between 1½ and 3 minutes work on the test is 0.879. There are decided indications that double time would not be useful in improving the record of those whose score was high in the first 1½ minutes.

A very excellent criterion of the efficiency of a test is its value in differentiating between officers and men. In general, a sample of officers would contain a larger percentage of men of high intelligence than a sample of enlisted men. The amount that a test differentiates the groups would indicate the value of the test. This test of practical judgment was the worst test in the whole series in differentiating officers from men. If we used this criterion alone there would be no possible excuse for retaining the test in the series. In differentiating officers from men, it was about twice as bad as the next to the poorest test (oral directions).

On the other hand, we need tests in alpha which are effective at the lower end of the scale, and we can set up as our criterion here the value of the test in differentiating between feeble-minded individuals and enlisted men. The alpha tests were given to the high grade feeble-minded population of two custodial institutions, and the results compared with a group of 300 English speaking enlisted men. Test 3 proved to be very much superior to any other test in the series in differentiating between feeble-minded individuals and enlisted men. This fact more than justifies the inclusion of Test 3 in the scale.

All of these facts are difficult to interpret. My own interpretation is that the sixteen items did not measure or grade "practical judgment" in any sense, but that the inclusion of at least one very obviously false and really quite silly alternative in each item acted as an effective pitfall for the

feeble-minded. At least we are sure that the actual experimental results are conclusive enough to dispose of any and all arm-chair criticisms.

Alpha Test 4.—Synonym-Antonym

Time limit : 1½ minutes

If one will review the experimental literature of intelligence testing, he will find the synonym-antonym or "opposites" test, used sometimes as a test of controlled association, sometimes as a test of vocabulary, sometimes as a test of intelligence, but uniformly with excellent results. Given a group with a knowledge of English and sufficient intelligence to understand the nature of the problem, the synonym-antonym test will give as good a differentiation between the bright and dull members of the group, rated by an outside criterion, as any other standard test now available.

The distribution of the scores in Test 4 is shown in Figure 5.

The most striking feature of the distribution is the large number of scores that were either zero or one (393). This large number of zero scores is due to three causes. First, the illiterate group could not attempt it. Second, the stupid and literate could not understand the instructions and could not make the kind of judgment demanded. Third, in the long run chance or random responses would give scores around zero, for in scoring all tests that were a 50-50 guess, the total score was the number of right responses *minus* the number of wrong responses. If a person underlined "*same*" for every item, his score would be 20 right minus 20 wrong, or zero. If he merely guessed, he would,

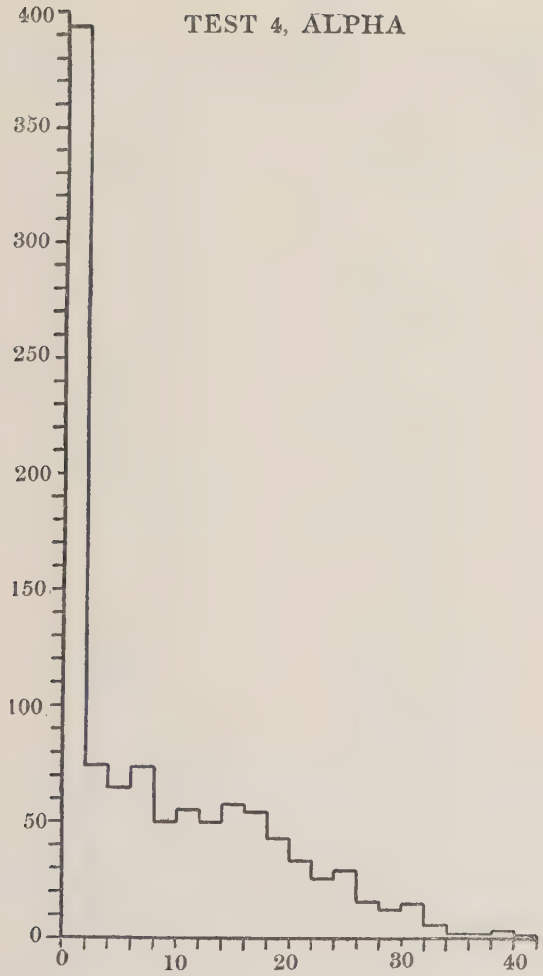

Figure 5. Distribution of scores of the Synonym-Antonym Test. (From p. 625, *Memoir XV.*)

If the two words of a pair mean the same or nearly the same, draw a line under *same*. If they mean the opposite, or nearly the opposite, draw a line under *opposite*. If you cannot be sure, guess. The two samples are already marked as they should be.

SAMPLES { good—bad ..... same—opposite
little—small ..... same—opposite

1	no—yes ..... same—opposite	1
2	day—night ..... same—opposite	2
3	go—leave ..... same—opposite	3
4	begin—commence ..... same—opposite	4
5	bitter—sweet ..... same—opposite	5
6	assume—suppose ..... same—opposite	6
7	command—obey ..... same—opposite	7
8	tease—plague ..... same—opposite	8
9	diligent—industrious ..... same—opposite	9
10	corrupt—honest ..... same—opposite	10
11	toward—from ..... same—opposite	11
12	masculine—feminine ..... same—opposite	12
13	complex—simple ..... same—opposite	13
14	sacred—hallowed ..... same—opposite	14
15	often—seldom ..... same—opposite	15
16	ancient—modern ..... same—opposite	16
17	enormous—gigantic ..... same—opposite	17
18	confer—grant ..... same—opposite	18
19	acquire—lose ..... same—opposite	19
20	compute—calculate ..... same—opposite	20
21	defile—purify ..... same—opposite	21
22	apprehensive—fearful ..... same—opposite	22
23	sterile—fertile ..... same—opposite	23
24	chasm—abyss ..... same—opposite	24
25	somber—gloomy ..... same—opposite	25
26	vestige—trace ..... same—opposite	26
27	vilify—praise ..... same—opposite	27
28	finite—limited ..... same—opposite	28
29	contradict—corroborate ..... same—opposite	29
30	immune—susceptible ..... same—opposite	30
31	credit—debit ..... same—opposite	31
32	assiduous—diligent ..... same—opposite	32
33	transient—permanent ..... same—opposite	33
34	palliate—mitigate ..... same—opposite	34
35	execrate—revile ..... same—opposite	35
36	extinct—extant ..... same—opposite	36
37	pertinent—relevant ..... same—opposite	37
38	synchronous—simultaneous ..... same—opposite	38
39	supercilious—disdainful ..... same—opposite	39
40	abstruse—recondite ..... same—opposite	40

Plate IV. Alpha Test 4: Synonym-Antonym (Form 8).

in the long run, guess half of the responses right and half wrong. Chance scores would then be zero (which includes all minus scores) and 1 or 2 on the positive side.

In general we may interpret Test 4 as a "high grade" test. It is too difficult to give any differentiation between low grade individuals, but it effectively grades the higher orders of intelligence. The time limit is not too short, for doubling the time only raises the average score from 10.50 to 12.60, and the correlation between regular and extended time is 0.940. It is one of the most effective tests in the scale for differentiating officers from enlisted men, and for differentiating feeble-minded from enlisted men. The only criticism is that it was too hard for a large number of people examined. Figure 5 really gives only about half of the normal distribution. If the test were so easy that the lower end of the scale could be extended to about −20, the distribution would become normal.

Alpha Test 5.—Disarranged Sentences

Time limit : 2 minutes

This test is an adaptation of a type of test which gives excellent results in the Binet-Simon scale. As it stands in alpha it is not a particularly good test. The distribution of scores shown in Figure 6 indicates a pile-up of zero scores due probably to the same three causes described as operating in Test 4. The test is fairly good in differentiating between officers and enlisted men, but for some reason or other it is the very worst test in the whole series in differentiating between feeble-minded and enlisted men. On the whole it is one of the poorest tests in our measuring scale.

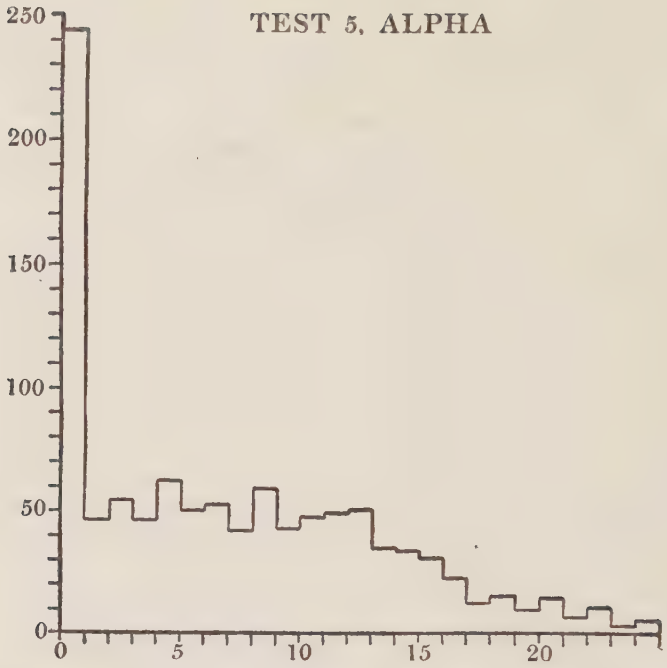

Figure 6. Distribution of scores of the Disarranged Sentence Test. (From p, 626, *Memoir XV.*)

The words A EATS COW GRASS in that order are mixed up and don't make a sentence; but they would make a sentence if put in the right order: A COW EATS GRASS, and this statement is true.

Again, the words HORSES FEATHERS HAVE ALL would make a sentence if put in the order ALL HORSES HAVE FEATHERS, but this statement is false.

Below are twenty-four mixed-up sentences. Some of them are true and some are false. When I say "go," take these sentences one at a time. Think what each would say if the words were straightened out, but don't write them yourself. Then, if what it would say is true, draw a line under the word "true"; if what it would say is false, draw a line under the word "false." If you can not be sure, guess. The two samples are already marked as they should be. Begin with No. 1 and work right down the page until time is called.

SAMPLES	a eats cow grass	true	false	
	horses feathers have all	true	false	
1	oranges yellow are	true	false	1
2	hear are with to ears	true	false	2
3	noise cannon never make a	true	false	3
4	trees in nests build birds	true	false	4
5	oil water not and will mix	true	false	5
6	bad are shots soldiers all	true	false	6
7	fuel wood are coal and for used	true	false	7
8	moon earth the only from feet twenty the is	true	false	8
9	to life water is necessary	true	false	9
10	are clothes all made cotton of	true	false	10
11	horses automobile an are than slower	true	false	11
12	tropics is in the produced rubber	true	false	12
13	leaves the trees in lose their fall	true	false	13
14	place pole is north comfortable a the	true	false	14
15	sand of made bread powder and is	true	false	15
16	sails is steamboat usually by propelled a	true	false	16
17	is the salty in water all lakes	true	false	17
18	usually judge can we actions man his by a	true	false	18
19	men misfortune have good never	true	false	19
20	tools valuable is for sharp making steel	true	false	20
21	due sometimes calamities are accident to	true	false	21
22	forget trifling friends grievances never	true	false	22
23	feeling is of painful exaltation the	true	false	23
24	begin a and apple acorn ant words with the	true	false	24

Plate V. Alpha Test 5: Disarranged Sentences (Form 8).

Alpha Test 6.—Number Series Completion

Time limit : 3 minutes

This test is the only one in alpha demanding a high order of intelligence almost entirely independently of the use of language. The greatest difficulty was experienced with the instructions for this test, when the first trial was made at the army camps. The preliminary forms contained only two rows of samples, and the instructions included the rather involved statement: "In the lines below, each number is gotten in a certain way from the numbers coming before it. Study out what this way is in each line and then write in the space left for it the number that should come next. The first two lines are already filled in as they should be." In the final alpha revision, four samples were included, and the instructions were simplified verbally and read very slowly. The instructions were given as follows: "Look at the first sample row of figures at the top of the page: 2, 4, 6, 8, 10, 12; the two numbers that should come next are, of course, 14, 16," etc., for each sample. Long pauses followed the reading of each sample.

The distribution of responses given in Figure 7 shows that the simplified instructions gave very good results, for although there were many zero scores in our experimental group of 1047 cases (244), there were probably no more zero scores than might have been expected when we consider that the mere understanding of what was wanted required considerable intelligence. On the whole the number series completion test proved to be entirely satisfactory.

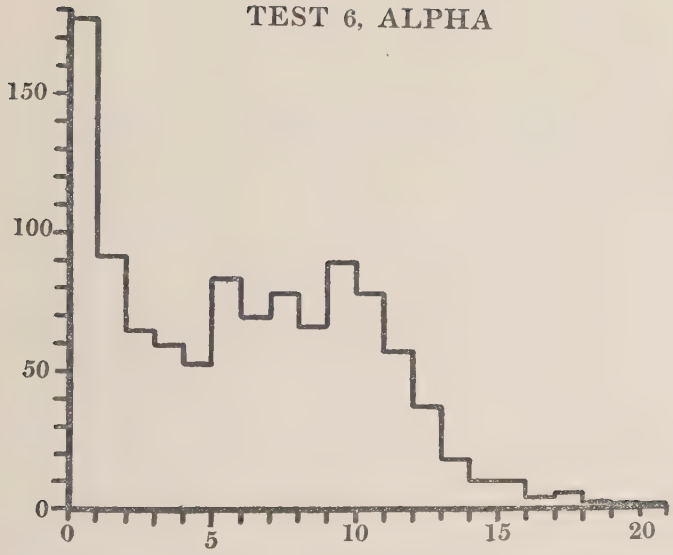

Figure 7. Distribution of scores of the Number Series Completion Test. (From p. 624, *Memoir XV.*)

SAMPLES	2	4	6	8	10	12	*14*	*16*
	9	8	7	6	5	4	*3*	*2*
	2	2	3	3	4	4	*5*	*5*
	1	7	2	7	3	7	*4*	*7*

Look at each row of numbers below, and on the two dotted lines write the two numbers that should come next.

3	4	5	6	7	8	.....	.....
8	7	6	5	4	3	.....	.....
10	15	20	25	30	35	.....	.....
9	9	7	7	5	5	.....	.....
3	6	9	12	15	18	.....	.....
8	1	6	1	4	1	.....	.....
5	9	13	17	21	25	.....	.....
8	9	12	13	16	17	.....	.....
27	27	23	23	19	19	.....	.....
1	2	4	8	16	32	.....	.....
19	16	14	11	9	6	.....	.....
11	13	12	14	13	15	.....	.....
2	3	5	8	12	17	.....	.....
18	14	17	13	16	12	.....	.....
29	28	26	23	19	14	.....	.....
20	17	15	14	11	9	.....	.....
81	27	9	3	1	$\frac{1}{3}$	.....	.....
1	4	9	16	25	36	.....	.....
16	17	15	18	14	19	.....	.....
3	6	8	16	18	36	.....	.....

Plate VI. Alpha Test 6: Number Series Completion (Form 8).

Alpha Test 7.—Analogies

Time limit : 3 minutes

The analogies test gave results on a par with the synonym-antonym test. The distribution of scores given in Figure 8 shows a large number of zero scores (284), but this number is not larger than might have been expected, considering the amount of intelligence necessary even to understand the nature of the task to be performed. This test again shows only a partial distribution. It it had been easier, it would probably have shown a symmetrical distribution.

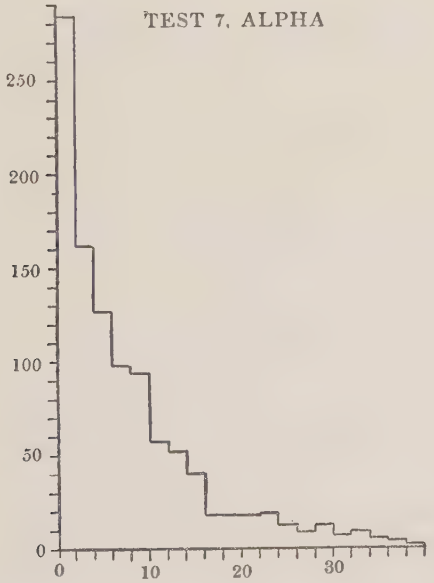

Figure 8. Distribution of scores of the Analogies Test. (From p. 625, *Memoir XV.*)

SAMPLES
sky—blue :: grass— table green warm big
fish—swims :: man— paper time walks girl
day—night :: white— red black clear pure

In each of the lines below, the first two words are related to each other in some way. What you are to do in each line is to see what the relation is between the first two words, and underline the word in heavy type that is related in the same way to the third word. Begin with No. 1 and mark as many sets as you can before time is called.

1	shoe—foot :: hat— kitten head knife penny	1
2	pup—dog :: lamb— red door sheep book	2
3	spring—summer :: autumn— winter warm harvest rise	3
4	devil—angel :: bad— mean disobedient defamed good	4
5	finger—hand :: toe— body foot skin nail	5
6	legs—frog :: wings— eat swim bird nest	6
7	chew—teeth :: smell— sweet stink odor nose	7
8	lion—roar :: dog— drive pony bark harness	8
9	cat—tiger :: dog— wolf bark bite snap	9
10	good—bad :: long— tall big snake short	10
11	giant—large :: dwarf— jungle small beard ugly	11
12	winter—season :: January— February day month Christmas	12
13	skating—winter :: swimming— diving floating hole summer	13
14	blonde—light :: brunette— dark hair brilliant blonde	14
15	love—friend :: hate— malice saint enemy dislike	15
16	egg—bird :: seed— grow plant crack germinate	16
17	dig—trench :: build— run house spade bullet	17
18	agree—quarrel :: friend— comrade need mother enemy	18
19	palace—king :: hut— peasant cottage farm city	19
20	cloud-burst—shower :: cyclone— bath breeze destroy West	20
21	Washington—Adams :: first— president second last Bryan	21
22	parents—command :: children— men shall women obey	22
23	diamond—rare :: iron— common silver ore steel	23
24	yes—affirmative :: no— think knowledge yes negative	24
25	hour—day :: day— night week hour noon	25
26	eye—head :: window— key floor room door	26
27	clothes—man :: hair— horse comb beard hat	27
28	draw—picture :: make— destroy table break hard	28
29	automobile—wagon :: motorcycle— ride speed bicycle car	29
30	granary—wheat :: library— read books paper chairs	30
31	Caucasian—English :: Mongolian— Chinese Indian negro yellow	31
32	Indiana—United States :: part— hair China Ohio whole	32
33	esteem—despise :: friends— Quakers enemies lovers men	33
34	abide—stay :: depart— come hence leave late	34
35	abundant—scarce :: cheap— buy costly bargain nasty	35
36	whale—large :: thunder— loud rain lightning kill	36
37	reward—hero :: punish— God everlasting pain traitor	37
38	music—soothing :: noise— hear distracting sound report	38
39	book—writer :: statue— sculptor liberty picture state	39
40	wound—pain :: health — sickness disease exhilaration doctor	40

Plate VII. Alpha Test 7: Analogies (Form 8).

In the construction of this test it was deliberately planned to make many items very difficult by introducing as a wrong alternative a word that was frequently associated with the key word. For instance, we find in the last 20 items the following easily associated pairs which would be wrong: (21) first-last, (22) no-yes, (25) day-night, (26) window-door, (27) hair-comb, (28) make-break, (29) motor-cycle-ride, (30) library-read, (32) part-hair, (35) cheap-buy, (36) thunder-lightning, (38) noise-hear, (39) statue-liberty, (40) health-sickness. The test therefore involves not only the selection of the right word, but the refusal to accept as the solution a word that is exceedingly attractive owing to frequent associations.

The analogies test is the most effective test in the entire series in differentiating officers from men. For some reason, not understood, it does not rank high in differentiating feeble-minded from enlisted men. The scores show a considerable average improvement with extension of time limit (8.60 to 12.46) but a correlation of 0.920 between three and six minutes work. On the whole it is an excellent test.

Alpha Test 8.—Information

Time limit : 4 minutes

The army information test has been criticised more than any other test, and it has undoubtedly received much more abuse than it deserves. From the standpoint of test construction it is satisfactory, for, aside from the zero scores probably due to illiteracy, the distribution as shown in Figure 9 is rather good.

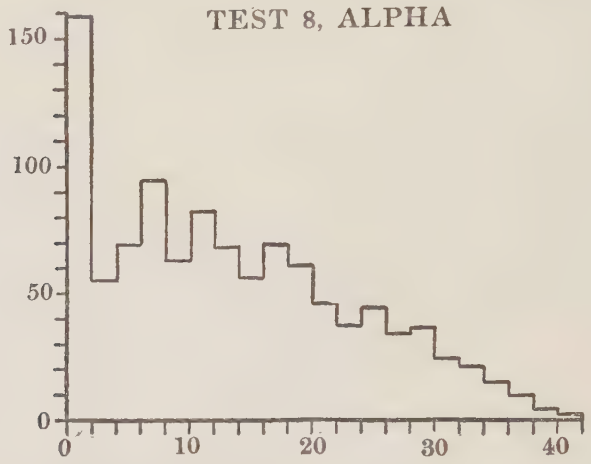

Figure 9. Distribution of scores of the Information Test. (From p. 626, *Memoir XV.*)

The most frequent charge made against the test is that a person could fail in certain items and still be intelligent. This is certainly true, and the criticism would be valid if anyone were expected to answer all the items, or if he were considered unintelligent if he failed. The average person answered less than 15 items correctly.

Notice the sample sentence:

People hear with the eyes <u>ears</u> nose mouth

The correct word is ears, because it makes the truest sentence.

In each of the sentences below you have four choices for the last word. Only one of them is correct. In each sentence draw a line under the one of these four words which makes the truest sentence. If you can not be sure, guess. The two samples are already marked as they should be.

SAMPLES { People hear with the eyes <u>ears</u> nose mouth
France is in <u>Europe</u> Asia Africa Australia

1	The apple grows on a shrub vine bush tree	1
2	Five hundred is played with rackets pins cards dice	2
3	The Percheron is a kind of goat horse cow sheep	3
4	The most prominent industry of Gloucester is fishing packing brewing automobiles	4
5	Sapphires are usually blue red green yellow	5
6	The Rhode Island Red is a kind of horse granite cattle fowl	6
7	Christie Mathewson is famous as a writer artist baseball player comedian	7
8	Revolvers are made by Swift & Co. Smith & Wesson W. L. Douglas B. T. Babbitt.	8
9	Carrie Nation is known as a singer temperance agitator suffragist nurse	9
10	"There's a reason" is an "ad" for a drink revolver flour cleanser	10
11	Artichoke is a kind of hay corn vegetable fodder	11
12	Chard is a fish lizard vegetable snake	12
13	Cornell University is at Ithaca Cambridge Annapolis New Haven	13
14	Buenos Aires is a city of Spain Brazil Portugal Argentina	14
15	Ivory is obtained from elephants mines oysters reefs	15
16	Alfred Noyes is famous as a painter poet musician sculptor	16
17	The armadillo is a kind of ornamental shrub animal musical instrument dagger	17
18	The tendon of Achilles is in the heel head shoulder abdomen	18
19	Crisco is a patent medicine disinfectant tooth-paste food product	19
20	An aspen is a machine fabric tree drink	20
21	The sabre is a kind of musket sword cannon pistol	21
22	The mimeograph is a kind of typewriter copying machine phonograph pencil	22
23	Maroon is a food fabric drink color	23
24	The clarionet is used in music stenography book-binding lithography	24
25	Denim is a dance food fabric drink	25
26	The author of "Huckleberry Finn" is Poe Mark Twain Stevenson Hawthorne	26
27	Faraday was most famous in literature war religion science	27
28	Air and gasolene are mixed in the accelerator carburetor gear case differential	28
29	The Brooklyn Nationals are called the Giants Orioles Superbas Indians	29
30	Pasteur is most famous in politics literature war science	30
31	Becky Sharp appears in Vanity Fair Romola The Christmas Carol Henry IV	31
32	The number of a Kaffir's legs is two four six eight	32
33	Habeas corpus is a term used in medicine law theology pedagogy	33
34	Ensilage is a term used in fishing athletics farming hunting	34
35	The forward pass is used in tennis hockey football golf	35
36	General Lee surrendered at Appomattox in 1812 1865 1886 1832	36
37	The watt is used in measuring wind power rainfall water power electricity	37
38	The Pierce Arrow car is made in Buffalo Detroit Toledo Flint	38
39	Napoleon defeated the Austrians at Friedland Wagram Waterloo Leipzig	39
40	An irregular four-sided figure is called a scholium triangle trapezium pentagon	40

Plate VIII. Alpha Test 8: Information (Form 8).

The test was devised to sample as many fields of information as it was possible to sample with 40 items. In general the five information tests in the five forms of alpha sampled similar fields. For instance, the advertising slogans which appear in the five forms are "Hasn't scratched yet" (Form 5), "The makings of a nation" (Form 6), "Eventually, why not now?" (Form 7), "There's a reason" (Form 8), and "The flavor lasts" (Form 9), while the Overland, Buick, Rolls-Royce, Pierce Arrow and Packard appear in each of the five forms.

Information tests vary considerably in construction. There is a great difference between asking for the date of Lee's surrender and asking a person to choose between the four dates, 1812, 1865, 1886 and 1832. And again, a person is merely asked to elect whether the bassoon, xylophone, cymbal, clarionet and piccolo, appearing in each of the five forms of alpha, should be used in music, stenography, book binding, or lithography. Approximately a third of the times test for vocabulary rather than information in the literal sense. If a person, for instance, knows what a Zulu, or a Korean, or a Hottentot, or a Kaffir or a Papuan is, he very obviously knows the number of his legs.

As a rule women object to the information test more than men because the test samples rather heavily the fields of sport, mechanical interests, etc. The chances are that this test would penalize women rather heavily, but as a general rule the results of comparing the two sexes on alpha as a whole at various colleges show very slight differences in favor of the men. The sex differences found are not large enough to be significant.

At Camp Lee a group of 164 captains and 200 enlisted men of the same general intelligence level (i. e. A and B) were examined. The test showing the greatest differentiation between the two groups was the information test. The

only other test showing a difference in favor of the captains was Test 4 (synonym-antonym test), while the test for arithmetical reasoning (Test 2) and practical judgment (Test 3) showed differences in favor of the A and B enlisted men. The differences were, of course, small, but the greatest difference was shown by the information test. In differentiating between officers and the general run of enlisted men, the information test was fairly effective, and it was very nearly as good as the arithmetical reasoning and synonym-antonym test in differentiating between feeble-minded individuals and enlisted men, in spite of the fact that the feeble-minded obtained a somewhat higher percentage of their total score from this test than did the enlisted men.

After weighing all the evidence, it would seem that we are justified in ignoring most of the arm-chair criticisms of this test and in accepting the experimental evidence tending to show that the test was a fairly good one. The assumption underlying the use of a test of this type is that the more intelligent person has a broader range of general information than an unintelligent person. Our evidence shows that this assumption is, in the main, correct.

SECTION II

EXAMINATION BETA

When we turn to examination beta, we meet an entirely different problem, that of testing the intelligence of wholly or partially illiterate persons who could not take alpha on account of their language handicap, of testing non-English speaking persons some of whom knew only the simplest commands in English, and low grade individuals who did not have sufficient intelligence to make a substantial score on alpha. At the time of the first try-out of the army tests in the fall of 1917 at four cantonments, Devens, Dix, Lee and Taylor, examination *a*, the fore-runner of alpha, was in use, and various types of individual examinations were being tried out, but there was no non-verbal group test. To meet this need, a preliminary try-out of fifteen tests was made early in 1918, and a final examination composed of seven tests was subsequently published and widely used throughout the country. In the following pages the seven beta tests are reproduced in Plates 9 to 15, the method of administering them is described, and the results from the experimental group of 1047 cases are presented.

Examination beta was given under the most rigid experimental conditions. The experimenter stood on a platform back of which was a large black-board on which small duplicates of the seven tests could be shown one at a time. A demonstrator, whose duties were to act out the test problems on the black-board, was an essential part of the experiment. The experimenter in pantomime showed the demonstrator what to do on the black-board, then, after his

performance was completed, orderlies throughout the room explained to the men that they were to go ahead with the test and do what the demonstrator had done. The orderlies' vocabulary was limited to "Yes," "No," "Sure," "Good," "Quick," "Hurry up," "How many?" "Same," "Do it," "Fix it." The experimenter used just as few words as possible, and acted out every spoken sentence by pointing, motioning, etc. The demonstrator never spoke. His duties consisted of doing before the group just what the group was expected to do with the examination blanks.

Beta Test 1.—Maze

Time limit : 2 minutes

The black-board was turned so that two sample mazes as shown in Figure 10 appeared. The experimenter traced

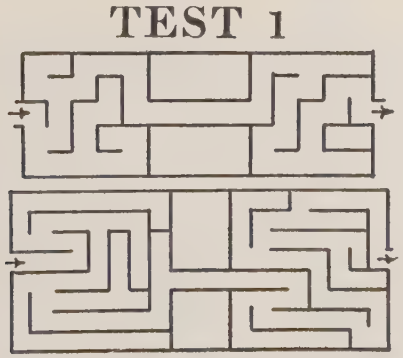

Figure 10. Black-board chart for demonstrating the Maze Test.

through the first maze on the black-board, and then motioned the demonstrator to go ahead. The demonstrator traced through the maze with crayon very slowly. The experimenter then traced through the second maze and motioned the demonstrator to go ahead. The demonstrator in tracing this maze made a mistake by crossing the line at the end of a blind alley, was corrected by the experimenter with vigorous shakes of the head and "no-no," and made to re-trace his path back to where he could start right again. The demonstrator then traced through the rest of

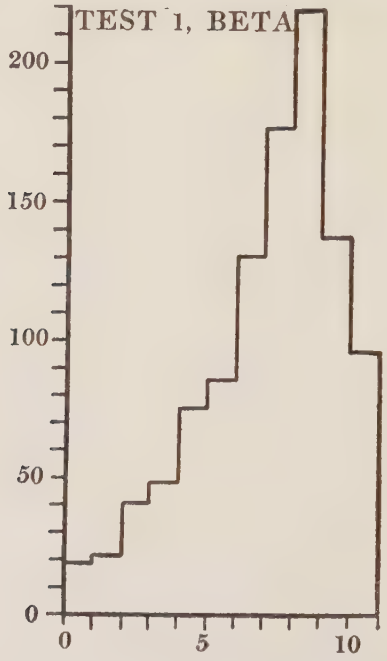

Figure 11. Distribution of scores of the Maze Test. (From p. 627, *Memoir XV.*)

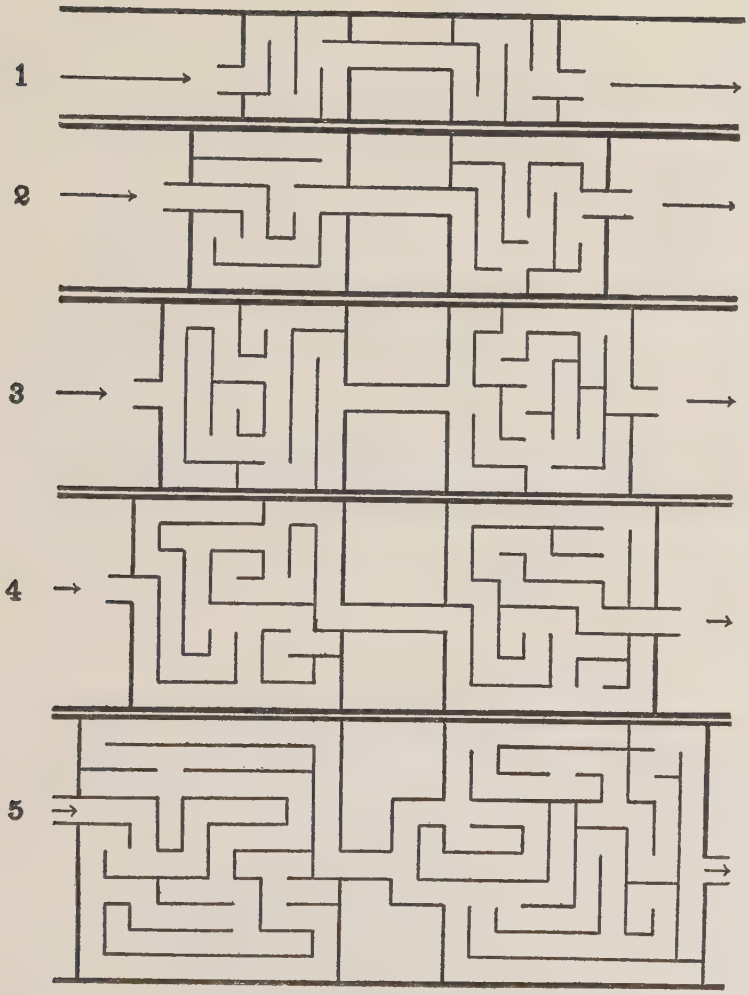

Plate IX. Beta Test 1: Maze.

the maze with great semblance of haste, stopping momentarily at each ambiguous point only. The experimenter then motioned to the group to do the same thing on their examination blanks. The experimenter and the orderlies walked about the room, motioning to the men who were not working, and saying, "Do it, do it, hurry up, quick."

The results of the maze test are shown in Figure 11. The difference between the distribution of scores of this test and the alpha tests is remarkable. In the first place, our large number of zero scores has disappeared—only 19 in our group of 1047 failed to make any score. In the second place, the test is entirely too easy, for it is apparent that the men in the upper end of the scale could have done more in the time allowed. The maze test was intentionally made easy in order to get everybody started. We have at last found a test in which practically everybody can do something. Aside from the language involved, every test in alpha is harder than this beta maze test, for no alpha test has less than 7% zero scores.

Beta Test 2.—Cube Analysis

Time limit : 2½ minutes

The black-board was turned to show a series of cubes like that in Figure 12. On a shelf was a real three cube

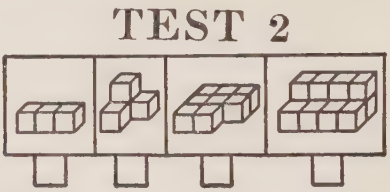

Figure 12. Blackboard chart for demonstrating the Cube Analysis Test.

model similar to the first one on the black-board. The experimenter pointed to the three cube picture on the black-board, then to the model on the shelf, then to the picture on the black-board, and asked, "How much?" The experimenter then counted aloud, putting up his fingers while counting, and encouraged the men to count with him. The experimenter then tapped each cube on the black-board and asked the demonstrator, "How much?" The demonstrator then went to the black-board, counted the cubes by pointing, and wrote the number 3 in the space below the illustration. A similar performance was enacted for the other three problems on the black-board, the models being shown and elaborately counted.

The distribution of scores of the cube analysis test is shown in Figure 13. Here we find a somewhat larger num-

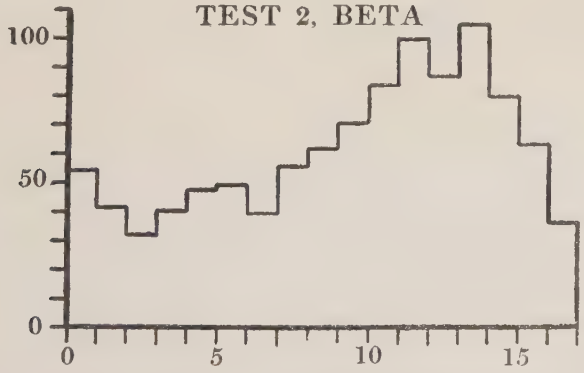

Figure 13. Distribution of scores of the Cube Analysis Test. (From p. 627, *Memoir XV.*)

ber of zero scores (54) than in the maze test, but a fairly good distribution in general. On the whole the test is easy.

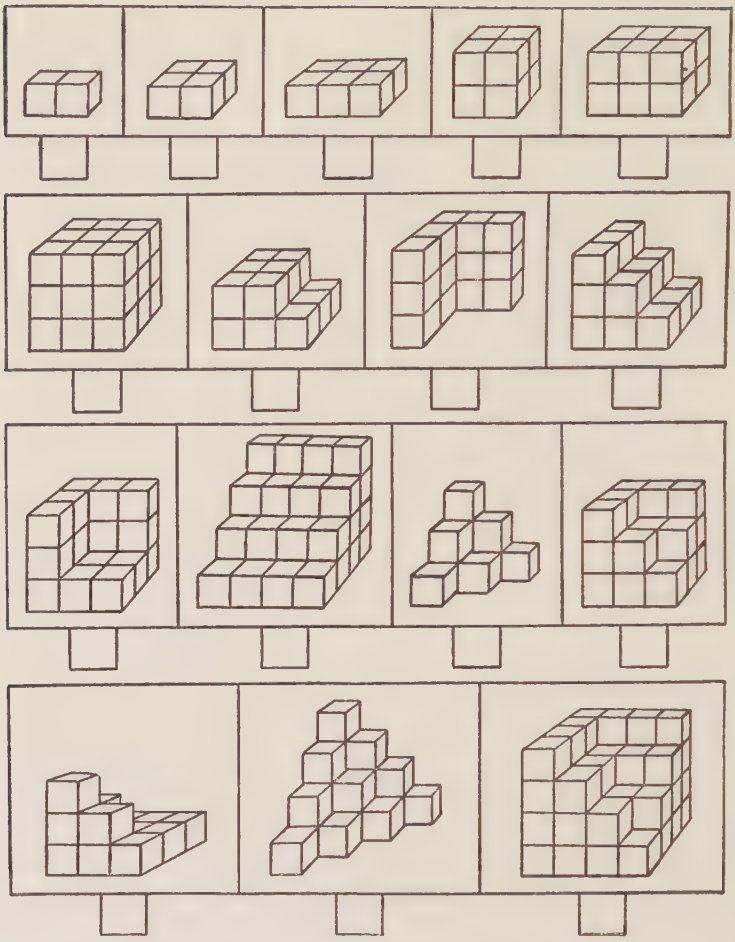

Plate X. Beta Test 2: Cube Analysis.

BETA TEST 3.—X-O SERIES

Time limit : 1¾ minutes

THE black-board was turned to show the chart reproduced in Figure 14. The experimenter traced with a pointer

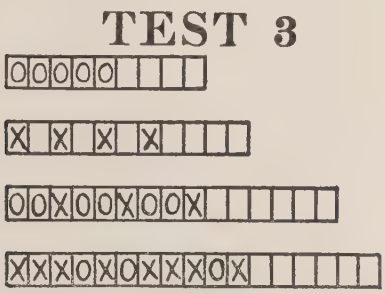

Figure 14. Black-board chart for demonstrating the X-O Series Test.

each "O" in the top chart, and then wrote (with his pointer) an imaginary "O" in the four remaining spaces. The demonstrator then filled in the four "O's" with crayon. The experimenter then traced the first "X" by tracing a semicircle above the chart and so on. The demonstrator then filled in every other space with an "X" following the experimenter's elaborate exercise. The demonstrator then worked out the remaining problems with the same ritual, following which the men were instructed to go ahead.

The distribution of scores of the X-O series test is shown in Figure 15.

The army writers state that: "Beta 3 defies interpretation." (p. 638.) We know that the test was devised to duplicate in pictorial form the number series completion test

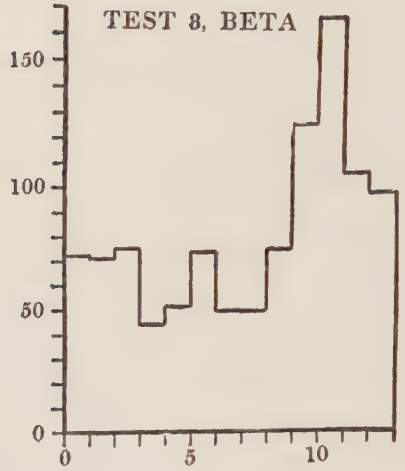

Figure 15. Distribution of scores of the X-O Series Test. (From p. 627, *Memoir XV.*)

in alpha. Aside from knowing the purpose of the test, we know very little about it. My own guess would be that the first five or six items were entirely too easy, and that if they had been disregarded in the scoring, as practice items, and six more items added, comparable in difficulty to the last six items, the distribution would have been satisfactory.

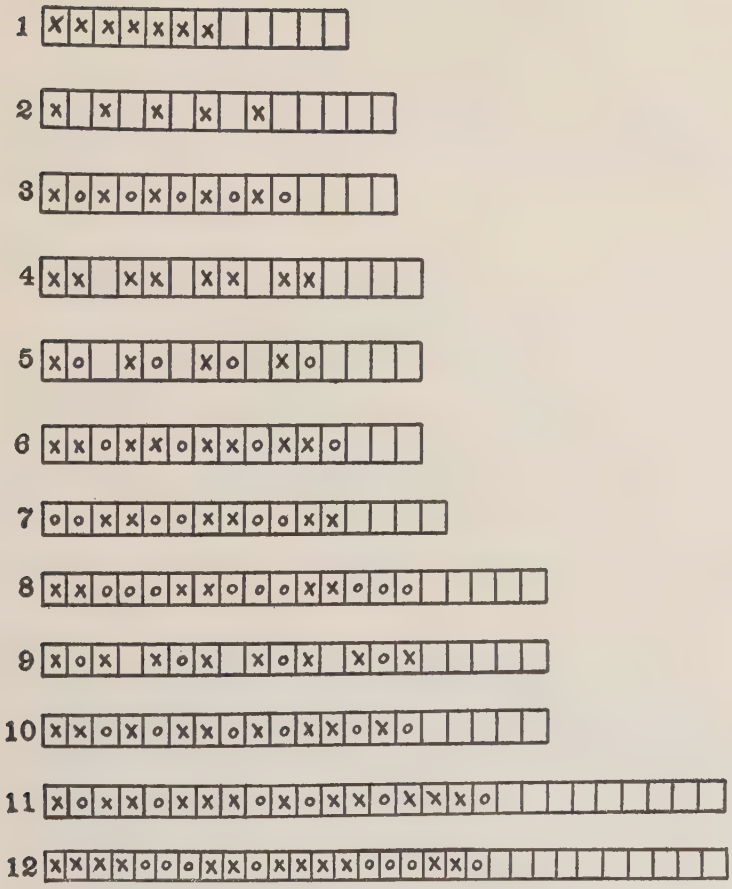

Plate XI. Beta Test 3: X-O Series.

Beta Test 4.—Digit-Symbol

Time limit : 2 minutes

The black-board was turned to the chart shown in Figure 16. The experimenter pointed to each number and then

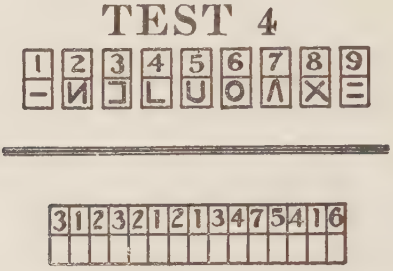

Figure 16. Black-board chart for demonstrating the Digit-Symbol Test.

to the symbol under it. The experimenter then pointed to the number 3 in the sample, then to the space below it, then to the number 3 in the key above, then to the symbol for 3, and finally traced the outline of the symbol for 3 in the proper space in the sample. This procedure was then repeated for the first five samples. The demonstrator then went to the black-board, and worked through the process of filling in the symbols under the figures, touching each figure and symbol in the key while he drew the proper symbol in the sample. The group then proceeded to fill in the symbols on the test blank.

The distribution of scores of the digit-symbol test is shown in Figure 17.

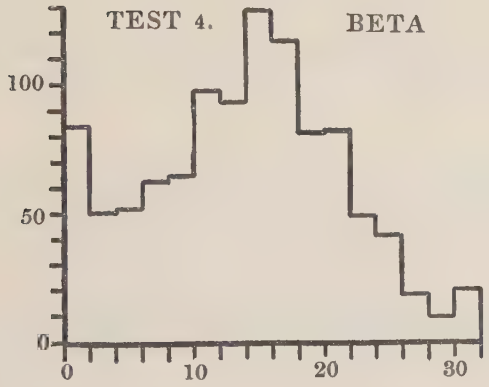

Figure 17. Distribution of scores of the Digit-Symbol Test. (From p. 628, *Memoir XV.*)

This type of distribution is the same as that given by the eight alpha tests, the zero scores representing failure to understand instructions, and the distribution being fairly regular. The digit-symbol test is a standard test, and the results in beta are entirely satisfactory.

1	2	3	4	5	6	7	8	9
−	И	⊐	L	∪	o	∧	X	=

1	3	1	2	1	3	2	1	4	2	3	5	2	9	1	4
2	6	3	1	5	4	2	7	6	3	8	7	2	9	5	4
3	6	3	7	2	8	1	9	5	8	4	7	3	6	9	5
4	1	9	2	8	3	7	4	6	5	9	4	8	5	7	6
5	9	3	8	6	4	1	5	7	2	6	2	4	8	1	3
6	4	9	5	1	7	5	2	6	9	3	7	8	4	1	8

Plate XII. Beta Test 4: Digit-Symbol.

Beta Test 5.—Number Checking

Time limit : 3 minutes

The black-board was turned to the chart shown in Figure 18. The experimenter pointed first to the 6 in the left

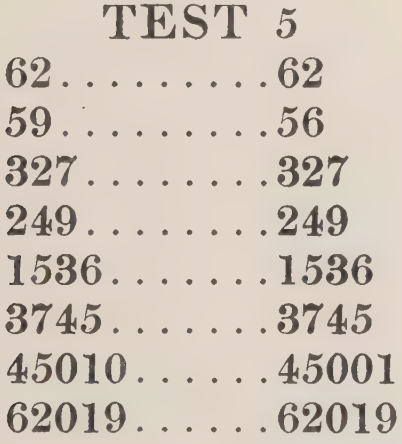

Figure 18. Black-board chart for demonstrating the Number Checking Test.

hand column, then to the 6 in the right hand column, then to the 2 in the left hand column and to the 2 in the right hand column, nodded his head, said, "Yes," and made an imaginary cross on the dotted line. The demonstrator then made an "X" on the line. The experimenter repeated the procedure for the second pair, but indicated clearly, by shaking his head and saying, "No," that the 9 and the 6 were not alike. The experimenter then repeated the procedure for three more sets, getting the men in the room to

say,"Yes," if the pairs were identical. The demonstrator then worked out the remaining items.

The results of the number checking test shown in Figure 19, give the same distribution characteristic of the

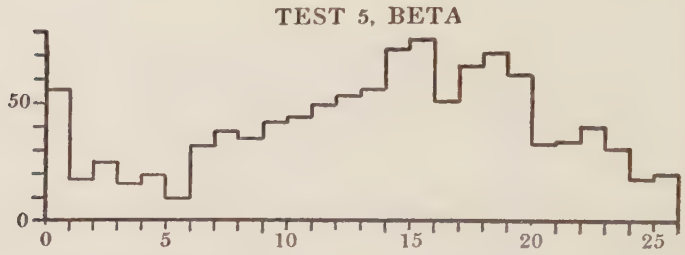

Figure 19. Distribution of scores of the Number Checking Test. (From p. 628, *Memoir XV.*)

alpha tests. The instructions were clear and the test was entirely satisfactory. This test is an adaptation of a standard test in use for many years.

650	650	10243586	10243586
041	044	659012534	6590211354
2579	2579	388172902	381872902
3281	3281	631027594	631027594
55190	55102	2499901354	2499901534
39190	39190	2261059310	2261659310
658049	650849	2911038227	2911038227
3295017	3290517	313377752	313377752
63015991	63019991	1012938567	1012938567
39007106	39007106	7166220988	7162220988
69931087	69931087	3177628449	3177682449
251004818	251004418	468672663	468672663
299056013	299056013	9104529003	9194529003
36015992	360155992	3484657120	3484657210
3910066482	391006482	8588172556	8581722556
8510273301	8510273301	3120166671	3120166671
263136990	263136996	7611348879	76111345879
451152903	451152903	26557239164	26557239164
3259016275	3295016725	8819002341	8819002341
582039144	582039144	6571018034	6571018034
61558529	61588529	38779762514	38779765214
211915883	219915883	39008126557	39008126657
670413822	670143822	75658100398	75658100398
17198591	17198591	41181900726	41181900726
36482991	36482991	6543920817	6543920871

Plate XIII. Beta Test 5: Number Checking.

Beta Test 6.—Picture Completion

Time limit : 3 minutes

The black-board was turned to the chart shown in Figure 20. The experimenter pointed to the hand and said,

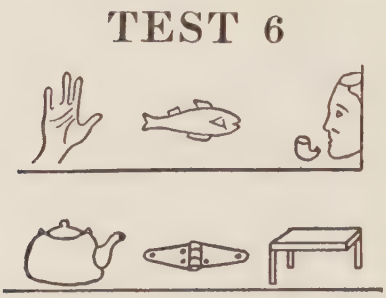

Figure 20. Black-board chart for demonstrating the Picture Completion Test.

"Fix it." The demonstrator looked puzzled. The experimenter pointed to the place where the finger was missing, and said, "Fix it; fix it." The demonstrator then drew the finger. The experimenter then pointed to the fish, and the place for the eye, and said,"Fix it." After the demonstrator had drawn in the eye, the experimenter pointed to each of the drawings and said, "Fix them all." After the demonstrator had worked out all the remaining drawings the group proceeded to complete the drawings in the beta blank.

The results of the pictorial completion test as given in Figure 21 show an excellent distribution of the same general type as the distribution of the eight alpha tests. The number of zero scores (12) on this test is smaller than that of any other test in the entire alpha-beta series. This is an excellent test.

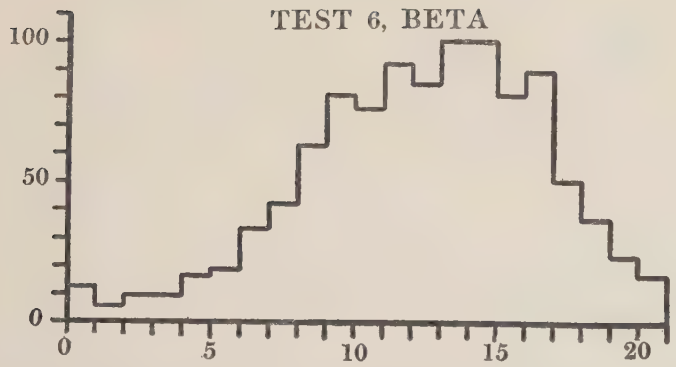

Figure 21. Distribution of scores of the Picture Completion Test. (From p. 628, *Memoir XV.*)

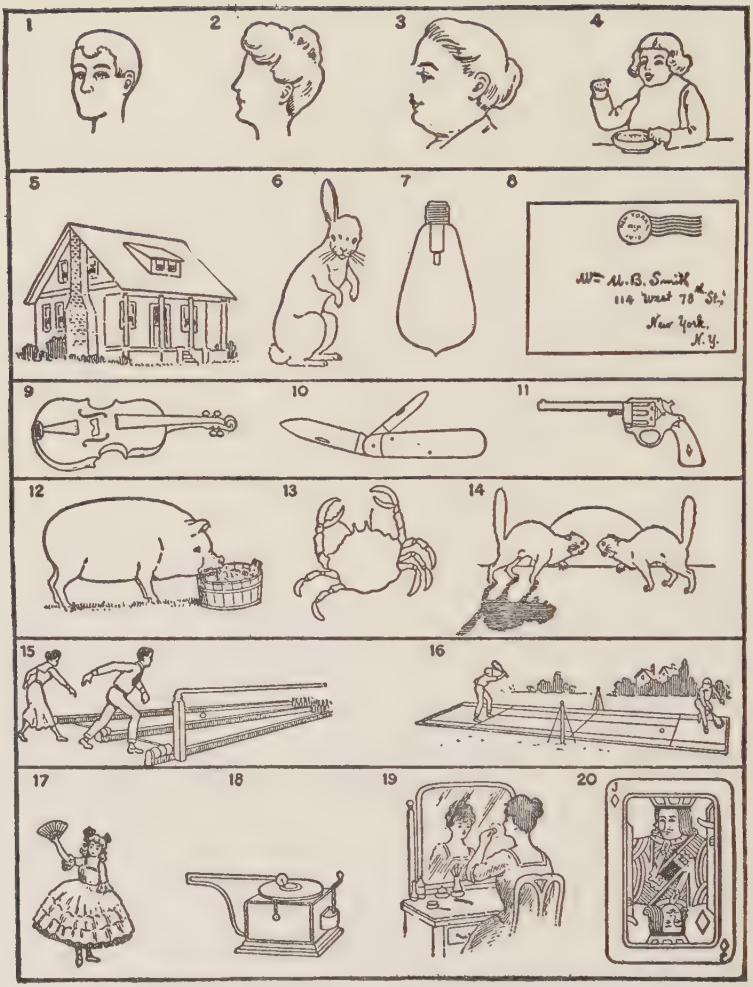

Plate XIV. Beta Test 6: Picture Completion.

Beta Test 7.—Geometrical Construction

Time limit : 2½ minutes

The black-board was turned to the chart shown in Figure 22. The experimenter pointed to the square on the

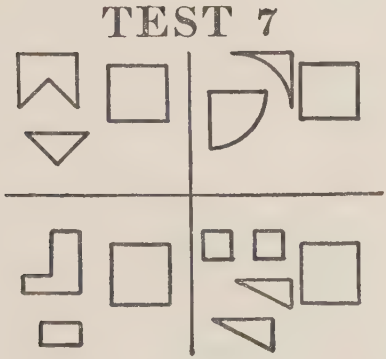

Figure 22. Black-board chart for demonstrating the Geometrical Construction Test.

black-board, and taking two pieces of cardboard the same size as the drawings at the left of the square, fitted them on the two drawings. He then fitted the two pieces of cardboard together on the square to show that they would fill it, and motioned to the demonstrator who drew, in the square, the lines indicating the manner in which the two pieces would fit. This procedure was repeated for the next two samples. The demonstrator then worked out the last sample.

The responses given in Figure 23 show, aside from the zero scores, a peculiar distribution, the form of which may be interpreted by assuming that the test was too hard in its

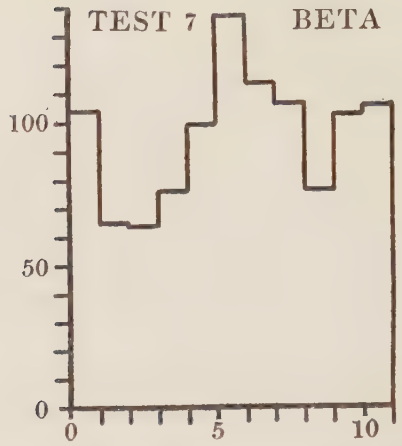

Figure 23. Distribution of scores of the Geometrical Construction Test. (From p. 628, *Memoir XV.*)

beginning and too easy at the end. We can picture the distribution of Test 7 as shown in Figure 23 as having come from the middle range of a more complete test such as Test 5, Figure 19. If we cut the distribution of Test 5 from 10 to 19, we can picture the range in which Test 7 was working. We may assume, then, that the inclusion of a few easier items and five or ten harder ones might have given a distribution similar to that of the alpha tests. The test is faulty because of the limitation of range at both ends.

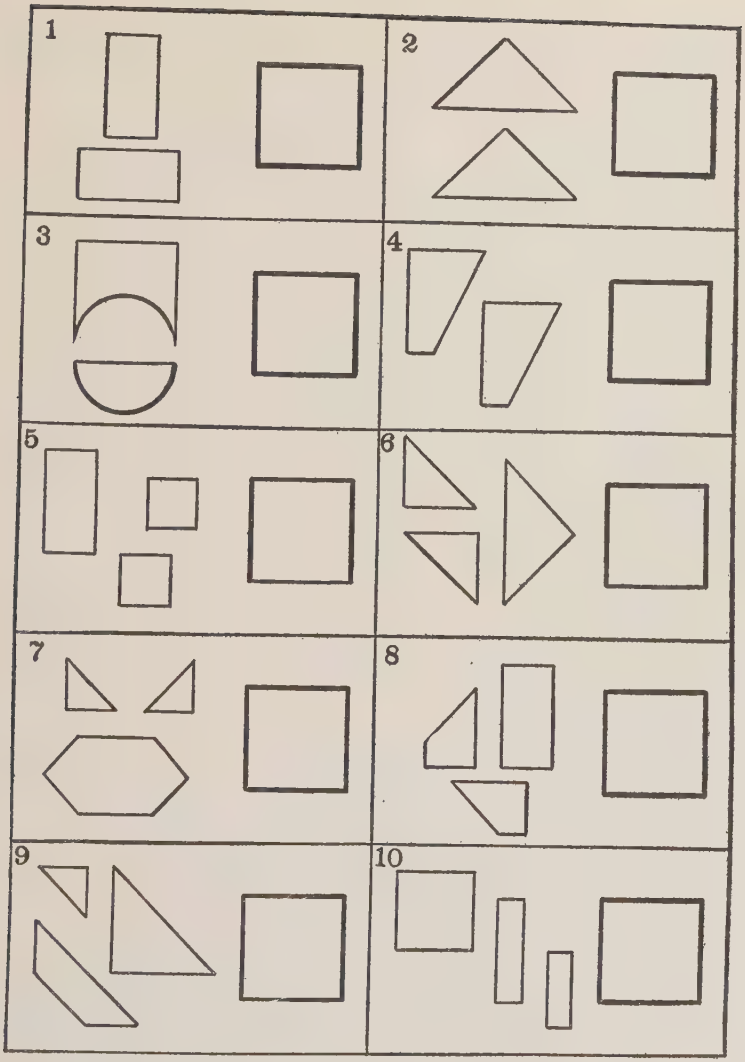

Plate XV. Beta Test 7: Geometrical Construction.

SECTION III

THE INDIVIDUAL EXAMINATIONS

The greatest contribution of the army psychologists to the development of mental tests was that of creating the two group tests, alpha and beta, that have been discussed. Methods of individual examining had been in existence for several years. The basic series of tests of the individual examination was the Stanford Revision of the Binet-Simon scale, which had become a standard measurement and needs no description here. Persons interested in this method should read Terman's[1] book on the Stanford scale. A method of abbreviating the Stanford-Binet scale was worked out in the army, and proved satisfactory.

The distribution of the scores in terms of the "mental ages" of the 653 men in the special experimental group of 1047 cases who took the Stanford-Binet examination was as follows:

"Mental Ages"	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
No. of cases	1	2	1	22	62	66	69	81	69	77	63	54	47	34	5

A rough inspection of these figures shows that they give us the Gaussian normal distribution. The results obtained from the Stanford-Binet examination may be taken as entirely reliable without question.

One difficulty in the popular interpretation of the results on the Stanford-Binet scale, and other scales constructed on the same principle, is the unfortunate use of the term "mental age," a term first used by Binet and subsequently

[1] L. M. Terman. *The Measurement of Intelligence.* Boston, 1916, pp. 362.

used in this country. The term "mental age" has no significance whatsoever aside from the particular scale from which it was derived. A person might have a "mental age" of 13 on the Stanford-Binet scale, of 11 on Goddard's translation of Binet's 1908 scale, of 12 on Goddard's 1911 scale, and so forth for every scale in use. The term "mental age" really means a score on a particular series of tests. Through rather general usage, the Stanford-Binet scale is being adopted in this country as a standard.

The Stanford-Binet scale was constructed out of some 90 different tests arranged for different age levels, six for each age level from 3 to 10, eight for 12, six for 14, six for 16 or "average adult," six for 18 or "superior adult," and sixteen alternative tests interspersed throughout the scale. A person obtains his total score or "mental age" by taking all the tests in perhaps four or five age level groups. For instance, if a person passed all the tests at the 9 year level, five out of the six at the 10 year level, four out of the eight at the 12 year level, two out of the six at the 14 year level, and failed all tests above 14, his total score or "mental age" would be 11½. In assigning a given test to any age level, all the tests were first tried out, and the positions of the tests juggled about so that the ten year old children tested had an average score of 10, the eleven year old children an average score of 11, etc.

When we say that a person has a "mental age" of eight on the Stanford-Binet scale, we do not mean that he has the mentality of a child of eight, but that he made a total score on that scale equal to that of the average eight year child tested in the particular group on which the scale was standardized. In all, about 1000 children, approximately equally distributed in the chronological ages from 5 to 14, formed the basis of the Stanford standardization. This standardization is a very excellent method of measuring

intelligence, and worked very well with our army adult group as shown by the score distribution of the 653 men in the special experimental group given above, but we should always regard the term "mental age" as a *score*, not as a diagnosis.

By correlating the alpha test with the Stanford-Binet scale, we can find the approximately equivalent score, or "mental age" for each possible alpha score. The operation resembles that of expressing values sterling in dollars. One frequently hears the statement that the army tests proved that the average citizen of this country has a mental age equivalent to that of a child of thirteen. Nothing could be more ridiculous. It is true that the average score of a sample of 93,955 soldiers representing the entire white draft, when translated into the Stanford-Binet scale, is 13.14. This means that the approximately equivalent score on the Stanford-Binet is 13.14. To say that the average citizen has a mentality of the child of 13 is putting the cart before the horse, for we are grading 93,955 people, and by inference the entire country, on a standard fixed by some 82 fourteen year old children who happened to be tested in California.

In addition to the 1000 children on whom the Stanford-Binet scale was standardized, the tests at the 16 year "average adult" level, and the 18 year "superior adult" level were standardized on 30 business men, 150 "migrating" unemployed men, 150 adolescent delinquents, and 50 high school students. It is thus seen that the Stanford-Binet standardization rests on a number of cases too small to upset the army standards based on 93,955 cases.

The term "mental age" is bad scientific slang for a total score. Psychologists are gradually abandoning the age standardization of tests. At the same time, publicists in various fields, although novices in psychology, are drawing

rather vicious conclusions from "mental age" findings. It is an unfortunate situation.

The methods used in creating and standardizing psychological tests are entirely empirical, and therefore rather hard to explain to the layman, who is familiar only with the "school teacher" type of examination. The school teacher writes an examination, and lets it stand as an absolute measure. The psychologist makes an examination, tries it out, and judges each individual member of the group as compared with the other members of the group. As more and more people are examined his standards of judgment become more reliable. In other words, his standards are those that he gets, not those that he thinks he ought to get. Therefore, instead of deploring the fact that the average person has a "mental age" of thirteen, we can simply say that the conversion of the results of the army test into the Stanford-Binet scale shows an average score of 13, and that this is the score to be expected from the average adult.

Another very common mis-statement prevalent concerning the army results is that they proved that 24.9% of the drafted men were illiterate. Among the men sent to examination beta would be found, first, English speaking illiterates, second, non-English speaking individuals, either literate or illiterate in their own tongue, third, defectives, and fourth, cases accidentally sent to the wrong examination. The method of selecting men for beta varied from camp to camp, and sometimes from week to week in the same camp. There was no established criterion of literacy, and no uniform method of selecting illiterates. In a group of 1,552,256 men examined, 386,196 or 24.9% were, for some reason or other, sent to beta. *The army definition of literacy as "ability to read and understand newspapers and write letters home" can not be identified with the fact of having*

been sent to beta. The statistics of the army examinations give us no accurate figures on the percentage of illiteracy.

The individual examination for illiterates and non-English speaking, the performance scale, was a composite scale, the tests of which were drawn from workers in various fields, particularly from H. A. Knox, who had worked on non-verbal performance tests at Ellis Island, R. Pintner and D. G. Paterson, who had developed a scale of performance tests, William Healy, H. H. Goddard, and other investigators. The performance examination was given sometimes as the long scale (8 or 10 tests) and sometimes as the short scale (5 tests). The short scale showed a correlation of 0.97 with the long scale, so that the reduction of the number of tests to save time in examining was entirely justified. A short-cut method was also used in giving the Stanford-Binet examination which was quite satisfactory (correlation 0.91). The short performance scale showed a correlation of 0.84 with the Stanford-Binet scale. The Yerkes-Bridges point scale which was sometimes used instead of the Stanford-Binet was also abbreviated, and the abbreviated point scale showed a correlation of 0.934 with the complete point scale. In general, the methods of individual examining in use were quite reliable, and so closely related to the Stanford-Binet scale that the results could be converted into Stanford-Binet scores without any appreciable source of error. In all calculations in this study, scores from the point scale and the performance scale examinations have been treated by converting into Stanford-Binet scores.

SECTION IV

RELIABILITY OF THE MEASURES

The reader, who has followed the discussion of the individual tests through the preceding pages, will be convinced that most of the tests used were satisfactory. In general, the eight alpha tests, the Stanford-Binet scale, and tests 4, 5, 6 and 7 of beta gave complete or limited distributions which approximated the Gaussian normal curve, Figure 24.

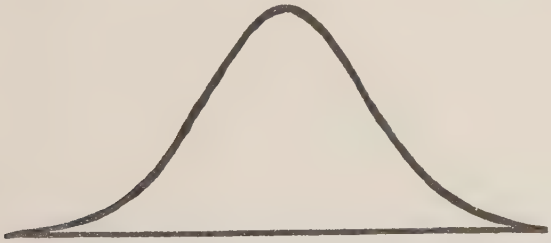

Figure 24. The normal distribution curve. A type of distribution given by all alpha tests, tests 4, 5, 6 and 7 of beta, and the Stanford-Binet scale.

Beta test 3 gave a distribution which could not be interpreted, while beta tests 1 and 2, being too easy, gave a skewed distribution of the approximate form of the curve shown in Figure 25.

It is not necessary here to enter into any lengthy discussion of the method of converting the results of these sixteen tests into the combined scale. The reader interested in the statistical methods used is referred to Chapter 2, Part 3 of Memoir XV (pp. 573–657). If the reader is satis-

fied that 13 of the 16 tests give distributions conforming in a general way to that shown in Figure 24, that is all that is necessary. The tests can very obviously be equated, and a combined scale constructed. By treating each of the eight tests of alpha, each of the seven tests of beta, and the Stanford-Binet test as different measuring scales, the combined scale was evolved, based on the inter-relations of these sixteen scales as shown by refined methods of correlation.

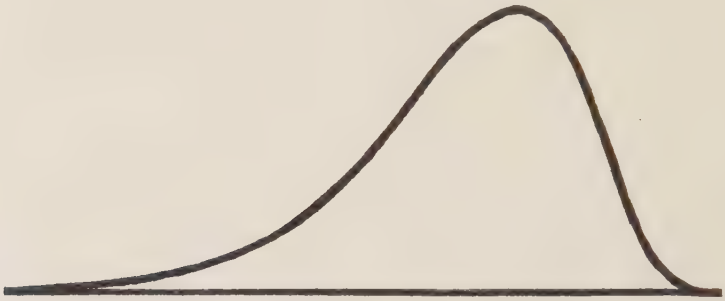

Figure 25. A skewed distribution curve. A type of distribution given by beta tests 1 and 2.

The army results are reported in tables showing the number of men scoring in certain class intervals, i. e. between 0 and 4, 5 and 9, 10 and 14, etc., up to the interval 205 to 212 on alpha; between 0 and 4, 5 and 9, 10 and 14, etc., up to 115–118 on beta. In the same way, the scores in other tests are reported in class intervals. The study of the 1047 cases showed how individuals falling in each of the class intervals were distributed on the theoretical combined scale, i. e., it was possible to find the combinations of tests from which individuals in each class interval would

obtain their scores. The combined scale was therefore built empirically on the results of the 1047 cases. Tables were constructed on this basis showing how individuals falling in each class interval of each of the three examinations should be redistributed on the combined scale. It is then possible to take a group which had been examined partly by alpha, partly by beta, and partly by the Stanford-Binet examination, and determine how that group would have scored on the combined scale if all individuals in the group had been given all three examinations. The combined scale is the most accurate method available for treating the data derived from the army examinations.

In this study the data from the principal sample have been re-figured on the combined scale by the method described on page 652 of Memoir XV:

"In each group the alpha distribution was distributed on the combined scale by the use of table 159, the beta distributions by table 162, and the Stanford-Binet mental age distribution by table 163. The performance scale distributions and the point scale distributions were handled in the following way: the performance distributions were first transformed into Stanford-Binet mental age distributions by the use of the regression formula:

$$\text{Mental Age (in years)} = \frac{0.50\ \text{Performance score} + 72}{12}$$

This formula was derived from the correlation of a sample of 350 cases who had both Stanford-Binet mental age ratings and performance scale ratings. The point scale distributions were transformed into Stanford-Binet mental age distributions by the use of the table in the examiner's guide, Part I, pages 195ff. These transformations only approximate the truth, but owing to the fact that the performance

and point scale cases constitute less than 3 per cent of any group handled it would take a considerable error in transformation seriously to affect the whole."

The conversion of the data of the principal sample into the combined scale reported in Memoir XV contains some inaccuracies. The statistical labor involved in the evolution of the combined scale was so great that the method was not available until the report was practically completed. The calculations were made by different individuals working under pressure, and errors were unavoidable. It has therefore been considered worth while to repeat these calculations at leisure, checking each operation carefully and carrying the analysis of some of the groups of the principal sample further than that reported in Memoir XV.

It is now necessary to review very briefly the results of checking the army mental tests against outside criteria. We might have a measuring scale, all elements of which gave perfect score distributions, and which were highly inter-correlated, but even then we would need outside criteria to prove that we were measuring intelligence. Enough material is already on hand to prove that the army tests were reliable measures of intelligence. In the following discussion we will cite several instances.

The best proof of the validity of the test series comes from a study of the relation between the intelligence ratings and education. The correlation of the combined scale with reported school grade was 0.75 (based on 653 cases from the special experimental group of 1047 men). The correlation between alpha scores and schooling for this group was 0.75, the eight tests of alpha separately compared with schooling all showing correlations between 0.60 and 0.74. The correlation with beta total scores and schooling for this same group was 0.67, and that between Stanford-Binet scores and schooling 0.65. These correlations show

a positive relationship between intelligence as measured by the various methods and years of schooling.

Very few people realize the severity of the elimination process that goes on from year to year in our schools and colleges. The study of the schooling of the native born white draft, as sampled by upward of 80,000 cases, showed the following startling facts: of every thousand native born recruits who entered the first grade, 970 remained in school till grade two, 940 till grade three, 905 till grade four, 830 till grade five, 735 till grade six, 630 till grade seven, and 490 till grade eight; 230 of them entered high school, 170 kept on till the end of the second year, 120 till the end of the third year, and 95 of the original thousand graduated from high school; 50 of these entered college, 40 kept on till the end of the second year of college, 20 till the end of the third year, and 10 graduated. It is, of course, impossible to determine how many of those that leave school leave on account of lack of pecuniary opportunity, or on account of lack of intelligence. It is ridiculous to assume that 1000 men in 1000 have sufficient intelligence to graduate from college, and equally absurd to assume that only 10 in 1000 have such a high intellectual endowment that they can graduate from college. But, inasmuch as we can not deny the intellectual elimination, we must expect a very high correlation between intelligence and schooling.

The army tests uniformly show officers superior to enlisted men. This is to be expected, for officers were selected for ability. Nevertheless, one may perhaps contend that the high scores of the officers were due to superior education and not to greater intelligence. Very nearly half of the officers were college graduates, and another quarter had begun but not completed a college course. The objection that the superior scores of officers were due to education rather than intelligence may be effectively answered by a

crucial test, which was made by the army investigators when they compared the alpha scores of 660 officers *who had never gone beyond the eighth grade in school* with the alpha scores of 13,943 native born recruits *all of whom had gone beyond the eighth grade.* The results of this comparison are reported (p. 779 of Memoir XV) as follows: "Every recruit in the recruit group has had more schooling than any officer in the officer group; the least educated recruit in the group has had a longer education than the best educated officer included. And the group of officers nevertheless makes a slightly higher record on examination alpha. It is evident then that the examination is measuring other qualities, in which officers stand above recruits, to a greater extent than it is measuring education." The distributions of the alpha scores of these two groups are shown in Figure 26. In general, the comparison of the army test scores with education indicates that the tests are genuine measures of intelligence.

The army investigators were, of course, called upon early in the war to prove that the tests they recommended were genuine tests of intelligence. For the assistance of army examiners and administrative officers having before them the problem of the assignment of men, a small pamphlet, *Army Mental Tests* (Washington, D. C., 1918, pp. 24) was prepared, presenting in graphic form the results of several different methods used in some of the camps for establishing the reliability of the tests by checking them against outside criteria. In the following pages, some of the charts from this booklet have been reproduced, and the method of interpreting the charts is described briefly. All of the methods reported use the letter grade classification (A, B, C, etc.), which is less accurate than the combined scale method used in this study, but they tend to show in

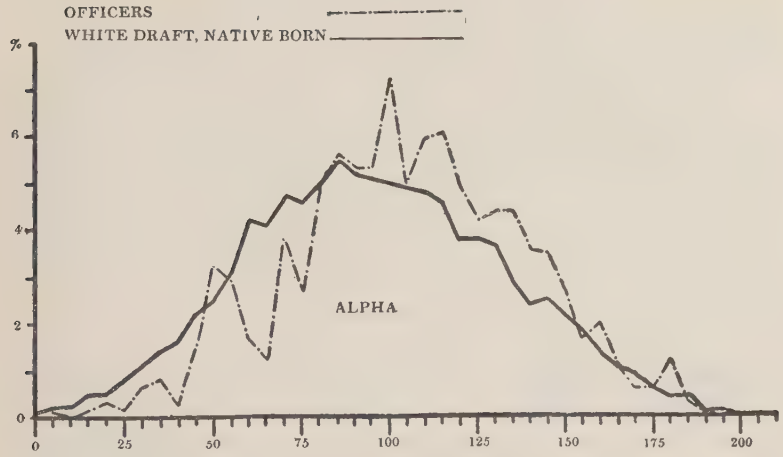

Figure 26. Examination alpha as independent of education. Comparison of alpha scores of officers of eighth grade schooling *or less* with alpha scores of native born white recruits of ninth grade schooling *or more.* "Although these groups overlap in schooling not at all, the officers make nevertheless slightly higher scores on alpha." (Quotation from p. 779, and figure from p. 778, *Memoir XV.*)

a general way the trend of the results, and that is all that is necessary in this case.

In the long run, we should expect a small positive correlation between intelligence and *rank*. Intelligence is by no means the sole determiner of military success, but since it is one element in the complex of abilities required, we would expect to find a general tendency toward high scores with higher ranks. Figure 27, which is reproduced from the

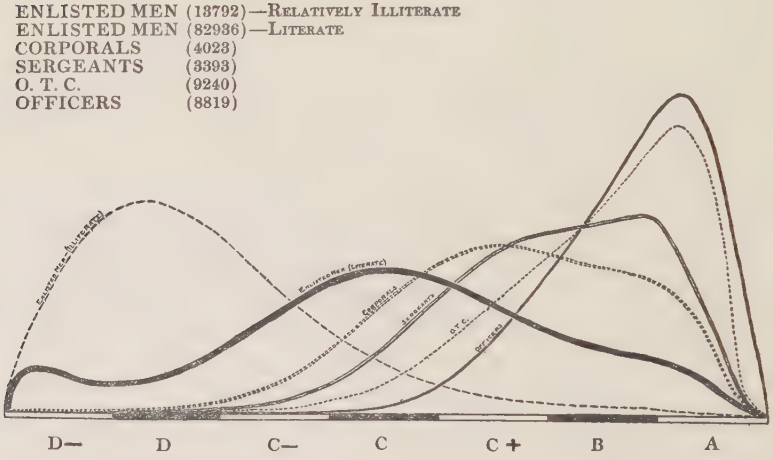

Figure 27. Distribution of intelligence scores according to rank. The officers are above the candidates in the Officers' Training Camps (O. T. C.), the candidate officers are above the sergeants, the sergeants above the corporals, and the corporals above the enlisted men. (From p. 8 of pamphlet, *Army Mental Tests.*)

booklet referred to, shows the distribution of scores of various ranks on the rough A, B, C scale. The officers form a group quite distinct from the general run of enlisted men, and they are also above the candidates for commissions in

the Officers' Training Camp (O. T. C.) group. The sergeants are above the corporals, and the corporals above the enlisted men.

The Officers' Training Camps give an additional check on the intelligence tests. In the schools examined, the candidates were recommended for a period of special training for commissions by the regimental organizations. The selection of the candidates was very rigid, then, in the first instance. Figure 28 shows roughly the results of applying the

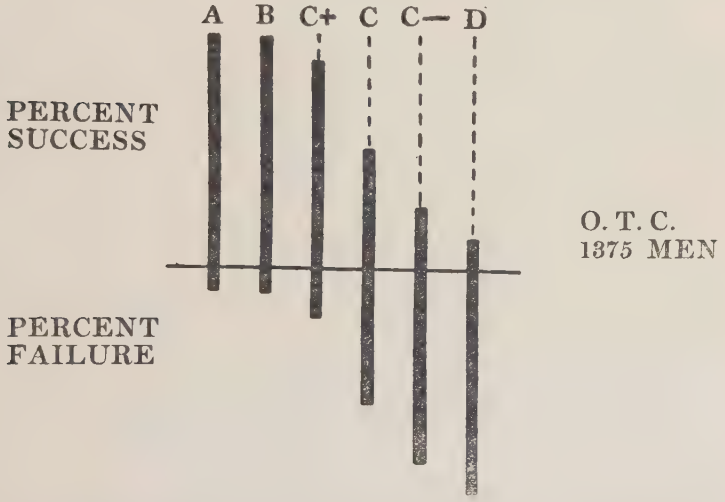

Figure 28. Success in Officers' Training Camps as predicted by examination alpha. Each vertical bar represents all (100%) of the candidates who tested A, B, C+, etc. All men above the horizontal line eventually received commissions, and all men below failed. 91⅓% of the men above C+ received commissions. 58⅓ of the men below C+ failed. (From p. 9 of pamphlet, *Army Mental Tests.*)

army tests to the training camp groups. Each solid vertical bar represents all the men of a given letter grade. There were, of course, more A and B men than C— and D men, but for purposes of comparison each letter group is treated as a whole (100%). All men above the horizontal line received commissions at the close of the Officers' Training Camp, and all the men below the line failed to receive commissions. Figure 28 shows clearly that about nine out of ten A and B men eventually received commissions, while for C— and D men the chances were very slight.

Figure 29 shows in a general way the manner in which groups selected by various outside criteria contributed to the upper end of the intelligence scale (A and B), and to the lower end of the scale (C—, D and E). We are already familiar with the differences between ranks shown in this figure. Sixty company commanders were asked to designate their ten "best" and ten "poorest" privates. The results of this comparison of the ten "best" and ten "poorest" privates also appears in Figure 29. The other two classifications, "men of low military value," and "unteachable men" represent a type of officers' rating. In general the test results check with officers' ratings independently made.

The results presented, showing the relation between rank and intelligence, and officers' ratings and intelligence, indicate clearly a certain positive relationship between tests and military success. Recognizing the fact that intelligence is only one factor tending to produce military success, we accept the results of checking the tests against military criteria as additional proof that the tests are genuine measures of intelligence.

A rough but rather interesting check of the army intelligence tests may be made by glancing at the scores of men classified by occupations. Figure 30 gives the range of the intelligence scores of the middle 50% of various occu-

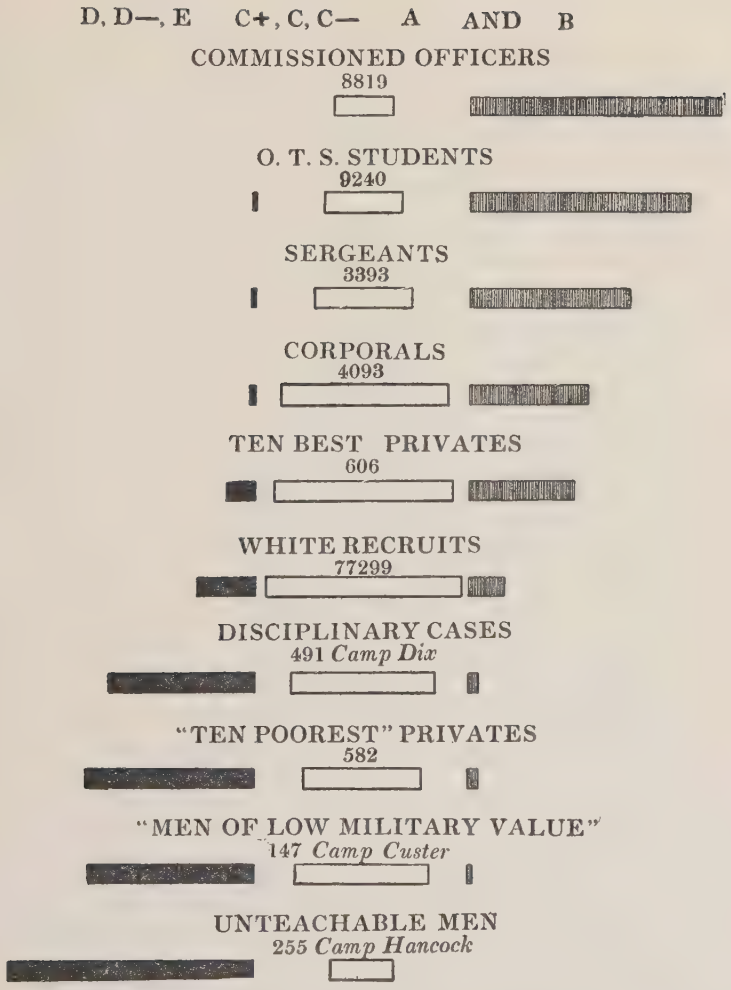

Figure 29. Comparison of army tests records with various independent criteria. The distributions of scores by ranks are shown in another way in Figure 27. The men were rated as "ten best," "ten poorest," "of low military value," and "unteachable" by their officers. The chart shows a close correspondence between the brief psychological examinations and officers' judgments made after weeks of observation. (From p. 10 of pamphlet, *Army Mental Tests.*)

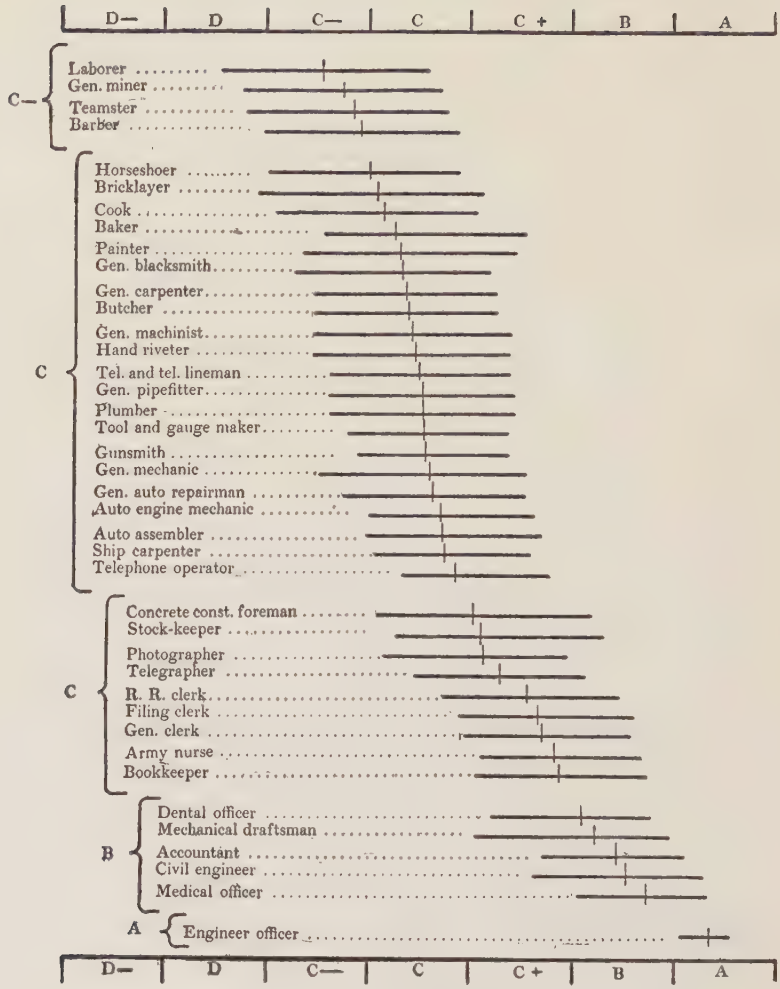

Figure 30. Success in civil occupations compared with army test records. The figure shows in a general way the correlation between intelligence as measured by the army tests and intelligence as indicated by position in civil life. (From p. 829, *Memoir XV.*)

pational groups, the position of the man half way up or down the scale being marked by a short vertical line. In some of the occupational groups, the number of cases is small, and the classification itself may be at fault in many instances, but the chart nevertheless shows a general tendency of the sort we should expect to find, for a process of intellectual selection occurs in industry which is just as rigid as that occurring in our public schools.

We have briefly inspected the different sorts of evidence from independent fields which indicate that the army tests were genuine measures of intelligence. Further discussion of this point is unnecessary. The army tests were not infallible, and mistakes in classifying men were undoubtedly made, but the tests were satisfactory rough measures. When used in comparing *groups* as the tests are in this study, their reliability is increased, for errors in measurement would tend to equalize in each group. We should expect the same percentage of error in classifying recruits born in Russia as we should recruits born in Sweden. Thus we use the tests as general measures of group tendencies, and as group measurements the tests have a sufficiently high degree of reliability to make positive conclusions possible.

In the foregoing pages the army tests have been described briefly, the method of treating the results from various examinations by the combined scale reviewed, certain misconceptions discussed, and a few bits of supporting evidence assembled. Persons interested in a further study of the tests should consult a little book by Yoakum and Yerkes[1], or, Memoir XV. We may now proceed to analyze American intelligence by treating the psychological examinations made in the army as a mental census of the population of this country.

[1]C. S. Yoakum and R. M. Yerkes. *Army Mental Tests.* New York, 1920, pp. 203

PART II

STATISTICAL ANALYSIS OF THE ARMY TEST RESULTS

PART II. SECTION I

THE PRINCIPAL SAMPLE

All results from the psychological examinations in the camps were sent to Washington. It was impracticable as well as undesirable to tabulate the results in the case of every man examined. An intelligent selection or sampling of cases will give results more nearly typical of the country at large, than the entire group tested, which would be unduly weighted for the more populous States, for camps giving the greatest number of examinations, particular draft quotas, etc. In order to obtain a sample for the white draft and the negro draft, cases were "randomly" (or better, impartially) selected in accordance with certain definite principles. The groups were as follows:

Group I: White draft, pro-rated, by States.... 41,278
Selected from 15 National Army camps, according to the State from which drafted, and according to the ratio of one recruit per thousand white male population.

Group II: White draft, additional, by States... 14,684
Additional selection of cases intended to bring the total representation from each State up to 1,000 cases.

Group III: White draft, additional, by camps... 40,392
Additional selection of cases intended to bring the entire sampling of the white draft up to approximately 100,000 cases.

Group IV: Negro draft, pro-rated, by States.... 19,992
Selected in the same manner as Group I.

Group V: Negro draft, additional, for Northern States..... 5,400
Chosen to represent the negro draft of the north.

Other groups were selected to meet other problems as follows:

Group VI: White officers..... 15,528
Selected proportionately to their occurrence in different arms of the service, with some additions to supplement the smaller arms, and the Medical Department.

Group VIII: White established organizations..... 24,205
Selected to provide comparison between enlisted men of various arms of the service.

Group X: Special experimental Group..... 1,047
Randomly selected individuals of the white draft born in English speaking countries, who were given both alpha and beta, and, where possible, the Stanford-Binet examination.

These groups selected as representative of the country at large were analyzed by the Hollerith system of mechanical sorting.

In this study we are concerned with:

Groups I, II and III representing the white draft,

Groups IV and V representing the negro draft,

Group VI representing the white officers, and Group X, the special experimental group.

SECTION II

ANALYSIS OF THE MAIN GROUPS OF THE PRINCIPAL SAMPLE

The tabulations in Memoir XV showing the distribution of scores on each type of examination of the white draft (Groups I, II, and III), the negro draft (Groups IV and V), and the white officers (Group VI) have been re-figured on the combined scale. The following tables were used:

For the white draft:

Alpha:	Table 183 (p. 666) for men who took alpha only	67,254
Beta:	Table 184 (p. 666) for men who took beta only, or alpha and beta	23,547
Individual:		
	Table 185 (p. 667) for men who took Stanford-Binet examination only, or following alpha, following beta, or following alpha and beta	1,246
	Table 186 (p. 667) for men who took point scale examination only, or following alpha, following beta, or following alpha and beta	689
	Table 187 (p. 668) for men who took performance scale examination only, or following alpha, following beta, or following alpha and beta	1,237
Total Cases		93,973

For the negro draft:

Alpha: Table 239 (p. 716) for men who took alpha only 8,429

Beta: Table 241 (p. 717) for men who took beta only, or alpha and beta 14,350

Individual:

Table 242 (p. 717) for men who took Stanford-Binet examination only, or following alpha, following beta, or following alpha and beta 403

Table 229 (p. 711) for men who took point scale examination only, or following alpha, following beta, or following alpha and beta 390

Table 228 (p. 710) for men who took performance scale examination only, or following alpha, following beta, or following alpha and beta 32

Total Cases 23,604

For White officers:

Alpha: Table 182 (p. 665) for all officers who took examination alpha only. 15,544

It will be remembered that examination beta was given to all men who had been selected as illiterate or non-English speaking before examination alpha was given, and also to those who took alpha and failed to make a reliable score. In the same way individual examinations were given to the lowest scoring cases in beta. Consequently, in figuring the results, if a man has taken both alpha and beta, his alpha score is disregarded and his beta score taken. In the same way, if he took both alpha and beta, and was then given an

individual examination, his alpha and beta scores are disregarded, and his score on the individual examination is taken as expressing the best measure of his intelligence. In other words we use alpha as our measure if alpha only was given, beta as our measure if beta alone, or alpha and beta were given, and the individual examination as our measure if it was given at all, on the assumption that the most reliable test of a man was the last one given.

The distribution of the scores of the white officers, white draft and negro draft is shown in Table 1. The first three columns are read in this manner: six officers and one recruit measured between 24.0 and 24.9 on the combined scale, one hundred and six officers and eighteen recruits between 23.0 and 23.9, etc. The first three columns show the actual distributions, i. e. six officers out of 15,543, one recruit out of 93,955, etc. The last three columns show each of these distributions arranged according to the number in each ten thousand who scored at each class interval of the combined scale. The last three columns read in this manner: four officers in ten thousand test between 24.0 and 24.9 on the combined scale; sixty-eight officers and two recruits in ten thousand score between 23.0 and 23.9 on the combined scale, etc. The "proportion in each ten thousand" may be read as a percentage by pointing off two decimal places.

Table 1 also shows the average score of the white officers, white draft and colored draft on the combined scale to be 18.84, 13.54 and 10.41 respectively. The standard deviation (S. D.) is also shown. An average has little significance without reference to a measure of variability of the series of measurements on which it is based. The conventionally accepted measure of variability is the standard deviation, which is derived by taking the square root of the average of the squares of the individual deviations from the aver-

Table No. 1

Distribution of the intelligence scores of the main groups of the principal sample on the combined scale.

COMBINED SCALE INTERVALS	ACTUAL DISTRIBUTION			PROPORTION IN EACH TEN THOUSAND		
	WHITE OFFICERS	WHITE DRAFT	NEGRO DRAFT	WHITE OFFICERS	WHITE DRAFT	NEGRO DRAFT
24.0–24.9	6	1	....	4	....	....
23.0–23.9	106	18	....	68	2	....
22.0–22.9	612	124	1	394	13	....
21.0–21.9	1648	444	7	1060	48	3
20.0–20.9	2522	1006	16	1623	107	7
19.0–19.9	2836	1804	35	1824	192	15
18.0–18.9	2698	2996	81	1736	319	34
17.0–17.9	2155	4687	172	1387	499	73
16.0–16.9	1454	6847	330	935	729	140
15.0–15.9	837	9328	600	538	993	254
14.0–14.9	412	12019	1031	265	1279	437
13.0–13.9	179	14659	1793	115	1560	760
12.0–12.9	60	14002	2572	39	1490	1090
11.0–11.9	14	9481	2951	9	1009	1251
10.0–10.9	3	6227	3187	2	662	1351
9.0–9.9	1	4433	3319	1	472	1406
8.0–8.9	....	2876	2891	....	306	1225
7.0–7.9	....	1683	2149	....	179	911
6.0–6.9	....	814	1315	....	87	557
5.0–5.9	....	334	684	....	36	290
4.0–4.9	....	122	302	....	13	128
3.0–3.9	....	37	112	....	4	48
2.0–2.9	....	11	38	....	1	16
1.0–1.9	....	2	10	....	....	4
No. cases.....	15543	93955	23596			
Average.....	18.84	13.54	10.41			
S. D.	2.10	2.92	2.79			

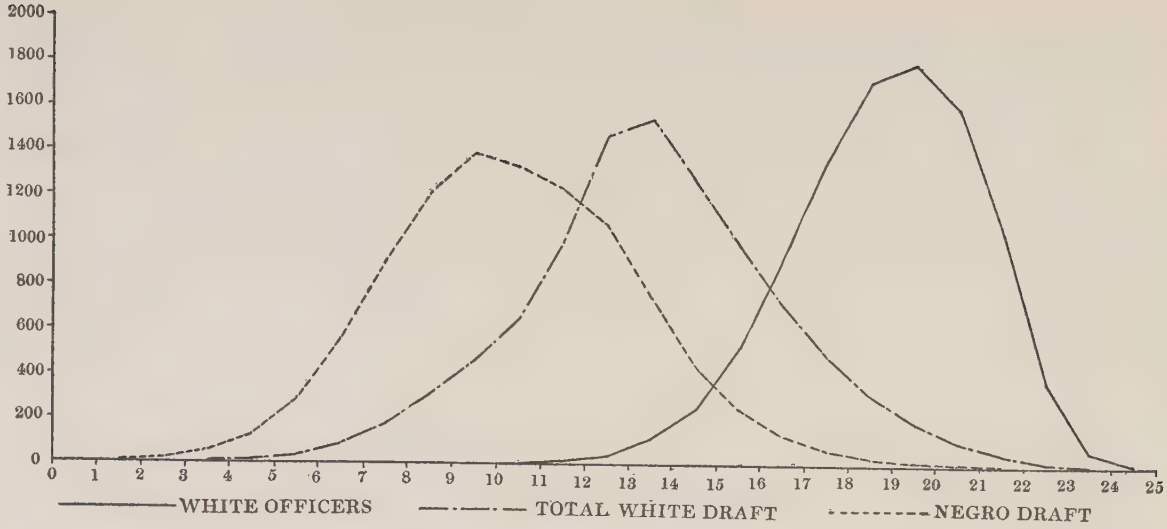

Figure 31. Distributions of scores of white officers, white draft and negro draft on the combined scale. The horizontal scale from 0 to 25 represents combined scale measurements, reading from left to right according to increasing intelligence. The vertical scale represents the number in each ten thousand. The three curves are drawn from the proportions given in the three right hand columns of Table 1. The differences in the intelligence of the three groups are very marked.

age. The size of the standard deviation indicates the variability of the group. In this case the white officers are more homogeneous (less variable) than the white draft and the negro draft (standard deviation of 2.10 compared with 2.92 and 2.79). Furthermore, the negro draft is more homogeneous than the white draft (2.79 compared with 2.92).

Figure 31 shows graphically the proportions given in the last columns of Table 1. The horizontal line shows the combined scale intervals and the vertical lines the number in each ten thousand. These curves show very clearly the distribution of intelligence in the three groups. In general, the distributions are similar in shape, but they differ markedly in their position on the scale of intelligence. The differences are very great. Of the officers, 98.87% are above the average of the white draft, and 99.97% are above the average of the negro draft. Of the white draft, 86.31% are above the average of the negro draft. Only 13.13% of the negro draft are above the average of the white draft. This method of figuring gives us some indication of the differences between the groups. If the distribution of intelligence in two groups were the same, 50% of either group would exceed the average of the other group. If the distributions were absolutely distinct, and there was no over-lapping, then 100% of one group would exceed the *highest man* in the other group. This last case would only occur if we compared very extreme groups (such as officers and idiots), and the conventional method is that of expressing the difference on a scale of 50%, i. e. the per cent. of one group above the *average* of the other group.

The most reliable method of determining the relationship between two groups is that of comparing the difference between the averages with the probable error of the difference. This method takes account of the variability of the original measures in each series, and also the reliability of

the measures as determined by both the variability and the number of cases. If the difference is -2 and the probable error of the difference is 1, the difference would be written -2 ± 1, meaning that the chances are even that it would not be less than -1, or more than -3. Differences which are not at least four times as great as the probable error of the difference are not conventionally accepted as significant.

Applying this method to the groups under consideration, we find the following differences:

Between white officers and white draft $-5.30 \pm .0131$.
(The difference is 405 times the probable error of the difference.)

Between white officers and negro draft $-8.43 \pm .0167$.
(The difference is 505 times the probable error of the difference.)

Between white draft and negro draft $-3.13 \pm .0138$.
(The difference is 227 times the probable error of the difference.)

SECTION III

ANALYSIS OF THE WHITE DRAFT INTO FOREIGN AND NATIVE BORN

THE next problem is that of breaking up the white draft into its constituent elements as far as possible. In Chapter 6, Part 3 of Memoir XV we find tables showing the distribution of the scores of 12,492 recruits who reported that they were born in foreign countries. The tables in Memoir XV give the scores of all but 800 who reported foreign birth. As all of the 12,492 reported cases were in Groups I, II and III, they may be deducted from those groups, leaving a remainder composed of (1) native born, and (2) foreign born who failed to report on their psychological examination blanks the fact that they were born in some country other than the United States. How large this latter group is we have no way of estimating. In the following tabulations the term "native born" is defined as all who stated that they were born in the United States, *plus* all who failed to record the country of their birth.

The original data giving the score distributions of the 12,492 foreign born were obtained from the following tables:

Alpha:	Table 207 (p. 693) for men who took alpha only	4,191
Beta:	Table 208 (p. 694) for men who took beta only, or alpha and beta	7,264
Individual:		
	Table 209 (p. 694) for men who took Stanford-Binet examination only, or following alpha, or following beta, or following alpha and beta	207

Table 210 (p. 596) for men who took performance scale examination only, or following alpha, following beta, or following alpha and beta	830
Total foreign born white draft	12,492

The actual distributions of the scores of the native born and foreign born are shown in the first two columns of Table 2, while the last two columns in this table give the proportion in ten thousand scoring at each class interval. These distributions are also shown graphically in Figure 32. Here again it is apparent that we have two groups that are markedly different. Of the native born 74.8% exceed the average of the foreign born. The difference between the native born and the foreign born is $-1.72 \pm .0186$, a difference that is $92\frac{1}{2}$ times the probable error of the difference.

Comparing the native born white draft with the negro draft shown in the preceding section, we find that 88.76% of the native born white draft exceed the average of the negro draft. The difference between these two groups is $-3.36 \pm .014$, a difference that is 240 times the probable error of the difference.

Comparing the foreign born with the negro draft, we find that 70.44% of the foreign born exceed the average of the negro draft. The difference between these two groups is $-1.64 \pm .0212$, a difference that is 77 times the probable error of the difference.

Table No. 2

Analysis of the white draft into foreign born and native born groups. Distribution of each group on the combined scale.

Combined scale intervals	Actual distribution		Proportion in each ten thousand	
	Native born white draft	Foreign born white draft	Native born white draft	Foreign born white draft
24.0–24.9	1	.....	.....	.....
23.0–23.9	18	1	2	1
22.0–22.9	120	4	15	3
21.0–21.9	428	16	52	13
20.0–20.9	971	35	119	21
19.0–19.9	1733	71	213	57
18.0–18.9	2850	147	349	117
17.0–17.9	4403	284	540	227
16.0–16.9	6345	502	779	402
15.0–15.9	8537	791	1048	633
14.0–14.9	10870	1148	1334	919
13.0–13.9	13066	1593	1604	1275
12.0–12.9	12220	1782	1500	1426
11.0–11.9	7885	1596	968	1277
10.0–10.9	4801	1425	589	1141
9.0–9.9	3178	1254	390	1004
8.0–8.9	1985	892	244	714
7.0–7.9	1153	530	142	424
6.0–6.9	556	259	68	207
5.0–5.9	228	106	28	94
4.0–4.9	83	39	12	31
3.0–3.9	25	12	3	10
2.0–2.9	7	4	1	3
1.0–1.9	2	1	.....	1
No. of cases....	81465	12492		
Average.....	13.77	12.05		
S. D.....	2.86	2.88		

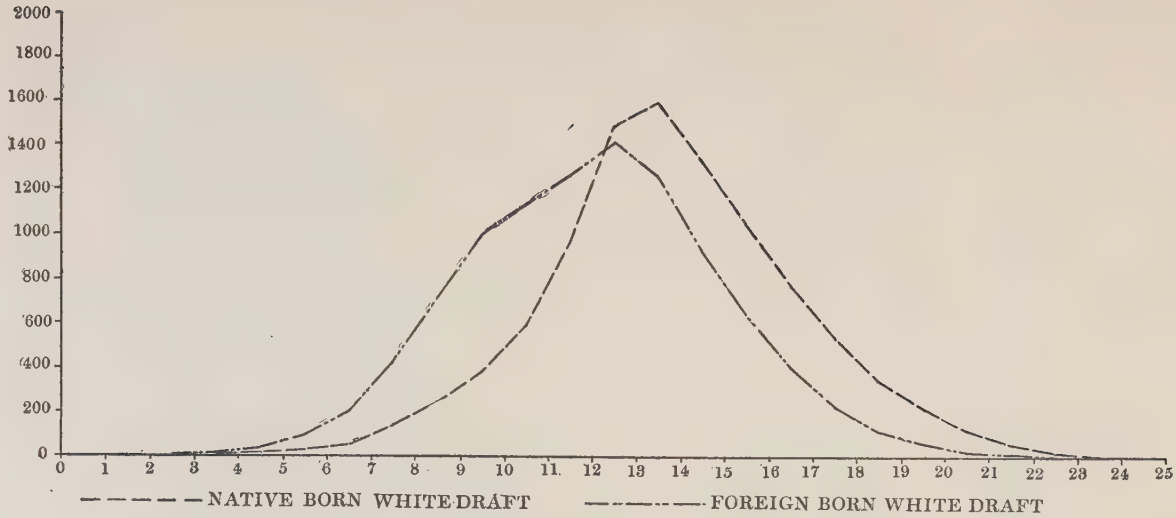

Figure 32. Distributions of scores of the native born and the foreign born white draft on the combined scale. Combined scale scores are shown on the horizontal scale, proportions in each ten thousand on the vertical scale. The curves are drawn from the proportions given in the two right hand columns of Table 2. These two curves are obtained by splitting the entire group representing the total white draft (given in Table 1, and shown in Figure 31) into its two components—the native born and foreign born. The curves show clearly that the foreign born are intellectually inferior to the native born. The foreign born group as a whole is, however, superior to the negro group.

SECTION IV

ANALYSIS OF THE FOREIGN BORN WHITE DRAFT INTO YEARS OF RESIDENCE GROUPS

The next problem is that of breaking up the foreign born white draft into its sub-groups, in order to discover the reason for the discrepancy between the main group of foreign born and the native born white draft. In Chapter 7, Part 3 of Memoir XV we find tables showing the scores of 11,295 foreign born (included in Groups I, II and III of the principal sample) classified according to the number of years of residence. The scores are tabulated for the following groups:

1st: a six year period..... 0 to 5 years of residence
2nd: a five year period.... 6 to 10 years of residence
3rd: a five year period....11 to 15 years of residence
4th: a five year period....16 to 20 years of residence
5th: a ten year period....Over 20 years of residence

As the age limit of the draft was 31, the last group would include cases who have been in the United States since childhood. It is probable that all of the 11,295 cases are included in the 12,492 cases shown in the preceding section. The discrepancy in numbers between the two groups is probably due to the fact that some foreign born reported the country of their birth, but failed to report the number of years they had been residents of the United States.

The scores reported in Chapter 7 have been re-figured on the combined scale. The following tables were used for the original data:

Alpha: Table 219 (p. 701) for men who took alpha only 3,619

Beta: Table 220 (p. 702) for men who took beta only, or alpha and beta 7,264

Individual:

Table 221 (p. 702) for men who took performance scale examination only, or following alpha, following beta, or following alpha and beta 802

Table 222 (p. 703) for men who took Stanford-Binet examination only, or following alpha, or following beta, or following alpha and beta 207

Total cases in all residence groups 11,295

The distributions of the scores of the five years of residence groups on the combined scale are shown in Table 3. This table shows a very remarkable fact, viz., a steady increase in the average scores with increasing years of residence, the averages being:

0 to 5 yrs.	6 to 10 yrs.	11 to 15 yrs.	16 to 20 yrs.	over 20 yrs.
11.41 (2.85)	11.74 (2.80)	12.47 (2.77)	13.55 (2.60)	13.82 (2.71)

From 0 to 20 years of residence, the average rises steadily and the variability becomes less and less.

Table 4 gives the difference between each group and every other group, together with the probable error of the difference, and the ratio of the difference to the probable error of the difference. All of the differences shown in Table 4 are significant except the difference between groups "16 to 20 yrs." and "over 20 yrs.," this difference (0.27) being only three times the probable error of the difference (± 0.0915).

Table No. 3

Analysis of the foreign born white draft by years of residence in the United States. Distribution of each residence group on the combined scale.

Combined scale intervals	Years of residence				
	0 to 5 years	6 to 10 years	11 to 15 years	16 to 20 years	over 20 years
24.0–24.9	....	....	0.1	....	....
23.0–23.9	....	0.1	0.5	....	....
22.0–22.9	0.6	0.6	0.8	0.5	0.5
21.0–21.9	2.6	2.9	2.6	1.6	3.6
20.0–20.9	5.8	8.2	6.3	3.6	7.5
19.0–19.9	14.4	18.6	13.1	8.3	12.1
18.0–18.9	27.6	37.6	27.9	18.0	24.7
17.0–17.9	55.5	72.0	50.6	36.7	41.1
16.0–16.9	104.5	141.8	83.1	61.2	66.2
15.0–15.9	172.4	240.7	131.3	87.3	86.1
14.0–14.9	261.2	355.2	198.7	106.1	106.3
13.0–13.9	374.4	506.2	282.0	128.6	128.6
12.0–12.9	444.8	597.0	309.1	124.5	113.7
11.0–11.9	457.5	572.1	245.6	82.6	69.0
10.0–10.9	471.0	533.0	189.8	47.4	42.9
9.0–9.9	453.3	479.2	152.3	29.5	29.5
8.0–8.9	342.6	347.2	102.1	18.1	18.2
7.0–7.9	212.9	209.6	58.5	9.9	9.2
6.0–6.9	106.8	102.4	27.0	4.5	3.4
5.0–5.9	44.7	41.7	10.5	1.7	1.0
4.0–4.9	16.5	14.9	3.6	0.6	0.3
3.0–3.9	5.0	4.4	1.1	0.2	0.1
2.0–2.9	1.5	1.3	0.3	0.1	....
1.0–1.9	0.4	0.3	0.1	....	....
No. of cases.....	3576	4287	1897	771	764
Average	11.41	11.74	12.47	13.55	13.82
S. D.....	2.85	2.80	2.77	2.60	2.71

Table No. 4

Comparison of the average scores on the combined scale of the five years of residence groups of the foreign born white draft.

YEARS OF RESIDENCE GROUPS		0 TO 5 YRS. AVE. 11.41 3576 CASES	6 TO 10 YRS. AVE. 11.74 4287 CASES	11 TO 15 YRS. AVE. 12.47 1897 CASES	16 TO 20 YRS. AVE. 13.55 771 CASES
6 to 10 yrs.	(Diff.	+0.33			
Ave. 11.74	(P. E. Diff.	±0.0431			
4287 cases	(Ratio	7.7			
11 to 15 yrs.	(Diff.	+1.06	+0.73		
Ave. 12.47	(P. E. Diff.	±0.0536	±0.0517		
1897 cases	(Ratio	19.8	14.1		
16 to 20 yrs.	(Diff.	+2.14	+1.81	+1.08	
Ave. 13.55	(P. E. Diff.	±0.0709	±0.0695	±0.0764	
771 cases	(Ratio	30.2	26.0	14.1	
Over 20 yrs.	(Diff.	+2.41	+2.08	+1.35	+0.27
Ave. 13.82	(P. E. Diff.	±0.0735	±0.0721	±0.0788	±0.0915
764 cases	(Ratio	32.8	28.8	17.1	3.0

Table No. 5

Comparison of the average scores on the combined scale of the native born white draft with the five years of residence groups of the foreign born white draft.

YEARS OF RESIDENCE GROUPS	NATIVE BORN WHITE DRAFT, 81465 CASES, AVE. 13.77 DIFFERENCE	PROBABLE ERROR OF THE DIFFERENCE	RATIO OF DIFFERENCE TO PROBABLE ERROR OF THE DIFFERENCE
0 to 5 yrs. Ave. 11.41 3576 cases	−2.36	±0.0104	226.9
6 to 10 yrs. Ave. 11.74 4287 cases	−2.03	±0.0296	68.6
11 to 15 yrs. Ave. 12.47 1897 cases	−1.30	±0.0434	30.0
16 to 20 yrs. Ave. 13.55 771 cases	−0.22	±0.0636	3.5
Over 20 yrs. Ave. 13.82 764 cases	+0.05	±0.0664	0.75

Table 5 shows the relationship of the five years of residence groups to the native born white draft. With increase in the time of residence, the differences between the native born and the foreign born become increasingly less significant. The difference between the native born and the "16 to 20 yrs." group of foreign born is slight (0.22) and is less than four times the probable error of the difference (± 0.0636). The foreign born group in this country over 20 years have an average score identical with the average score of the native born, the actual difference (0.05) being smaller than the probable error of the difference (± 0.0664).

Figure 33 shows graphically the relationships between the averages of the years of residence groups and the native draft. In this graph, the horizontal line represents increasing length of residence and the vertical line represents increase in the average score on the combined scale.

This very remarkable fact of increase in the intelligence score with years of residence was commented on by the army authors in Memoir XV as follows:

"It is not possible to state whether the difference is caused by the better adaptation of the more thoroughly Americanized group to the situation of the examination or whether some other factor is operative. It might be, for instance, that the more intelligent immigrants succeed and therefore remain in this country, but this suggestion is weakened by the fact that so many successful immigrants do return to Europe. At best we can but leave for future decision the question as to whether the differences represent a real difference of intelligence or an artifact of the method of examination." (p. 704.)

If our results reflect another factor independent of intelligence, which might be designated "the better adaptation of the more thoroughly Americanized group to the situation of the examination," we have no means of con-

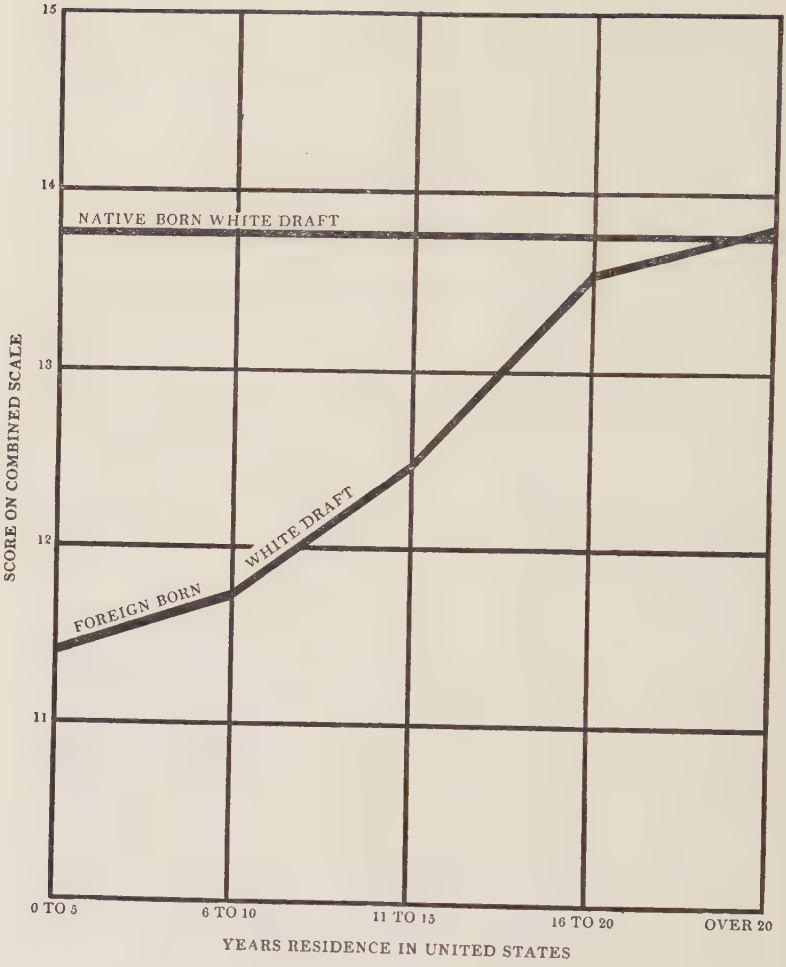
15
14
13
12
11
NATIVE BORN WHITE DRAFT
WHITE DRAFT
FOREIGN BORN
SCORE ON COMBINED SCALE
0 TO 5
6 TO 10
11 TO 15
16 TO 20
OVER 20
YEARS RESIDENCE IN UNITED STATES

Figure 33. Apparently increasing average intelligence with increasing years of residence. The horizontal scale reads from left to right according to increasing length of residence. The vertical scale represents score on the combined scale. For purposes of comparison, the position of the native born white draft on the combined scale is shown by a continuous line. The group of immigrants who have been in this country from 16 to 20 years have an average intelligence almost as high as that of the native born, while immigrants in this country over 20 years test the same as native born.

trolling this factor. Ultimately, the validity of our conclusions from this study rests on the validity of alpha, beta, and the individual examinations. It is sometimes stated that the examining methods stressed too much the hurry-up attitude frequently called typically American. The adjustment to test conditions is a part of the intelligence test. We have, of course, no other measure of adjustment aside from the total score on the examinations given. If the tests used included some mysterious type of situation that was "typically American," we are indeed fortunate, for this is America, and the purpose of our inquiry is that of obtaining a measure of the character of our immigration. Inability to respond to a "typically American" situation is obviously an undesirable trait.

For our purposes then we will accept the definition of intelligence given on page 573 of Memoir XV, viz., "by 'intelligence' we mean the ability that manifests itself quantitatively in a set of consistent scores in all of the types of examination upon which our data are based." We are forced to include the "adjustment to test conditions" in our definition of intelligence. And we hope, probably in the teeth of the facts, that the adjustment to test conditions involved a situation that was "typically American."

The hypothesis that the more intelligent immigrants remain in this country while the more stupid ones go home, which was offered by the army authors to account for the increase of intelligence scores with increasing years of residence, can not be checked from the data available in this study, and the emigration statistics give us little help. Table 6 shows the ratio between emigrant aliens and immigrant aliens from each country from 1908 to 1917. No figures for emigrant aliens are available prior to 1908.

Table 6 shows that since 1908 a very considerable number (approximately one third) of immigrants have event-

ually returned to their native countries. If the selection factor mentioned were operating, and in the long run the departing group contained more persons of lower intellectual capacity than the remaining group, an intelligence measurement would show an increase in the *direction* shown in Table 3, but the departing third would have to be very heavily weighted with low grade individuals to make any considerable difference in the distribution of the remaining two thirds.

Under the conditions of this study then, the hypothesis that the more intelligent immigrants succeed and therefore remain in this country must remain in the realm of speculation, as it can not be demonstrated as a fact. For our purposes the converse hypothesis, that the successful immigrants save their earnings and return to Europe to live in comfort for the rest of their lives, is equally cogent. Either hypothesis is a legitimate speculation. We must agree with the army authors that the first hypothesis is weakened by the fact that so many successful immigrants do return to Europe. The first hypothesis is also weakened by the fact that if the more inferior individuals left, the distribution of the intelligence of the departing third would have to be very markedly skewed at the lower end of the scale. The distribution curve of the departing third would be skewed to such an extent that only 10% or 15% of this group would exceed the median of the remaining two thirds. We must conclude that the selection factor mentioned might produce a slight change in the *direction* noted, but that it is highly improbable that any such factor could produce a change in the *amount* observed.

The important problem which we are facing is that of interpreting the fact of increase of intelligence test scores with increasing years of residence. Does our curve in Figure 33 represent the growth of intelligence with increasing

Table No. 6

Per cent. that emigration was of immigration for fifteen countries since 1908. The figures were obtained by dividing the number of emigrant aliens departed by the number of immigrant aliens admitted. The result is the ratio of emigration to immigration. The ratio 100 would mean that the number of alien immigrants admitted equalled the number of alien emigrants departed.

	AUSTRIA	BELGIUM	CANADA	DENMARK	ENGLAND	GERMANY	GREECE
1908.....	78.2	20.5	6.8	13.9	10.6	21.0	28.5
1909.....	29.1	11.7	58.6	10.5	9.4	19.2	40.0
1910.....	18.3	12.1	60.4	22.5	9.7	19.9	31.4
1911.....	54.2	17.8	86.9	6.2	10.3	18.8	35.7
1912.....	49.5	26.5	60.0	10.7	16.6	20.8	54.4
5 Yr. Period.....	42.8	17.5	57.9	12.8	11.3	19.9	37.2
1913.....	23.0	10.8	63.6	9.4	13.7	13.8	134.0
1914.....	27.0	19.9	37.0	10.0	20.2	14.4	31.0
1915.....	74.7	13.9	28.3	12.4	35.8	18.2	77.8
1916.....	16.1	2.4	15.5	15.4	39.8	15.3	17.8
1917.....	18.9	3.8	18.1	17.8	33.5	16.9	8.5
5 Yr. Period.....	26.7	13.7	30.4	12.0	23.6	14.6	47.7
10 Yr. Period 1908–1917.....	36.8	15.9	40.3	12.5	15.7	18.0	42.8

HOLLAND	IRELAND	ITALY	NORWAY	RUSSIA	SCOTLAND	SWEDEN	TURKEY	TOTAL ALL COUNTRIES
5.6	6.6	129.9	18.3	24.1	11.1	20.1	15.4	50.5
6.6	5.5	45.4	9.8	16.4	6.0	8.0	17.6	30.1
6.1	5.9	24.3	5.9	9.3	5.5	4.2	9.2	19.4
5.5	6.8	39.8	10.0	17.0	8.1	7.8	22.5	33.6
8.5	11.9	68.9	26.6	21.3	15.1	19.6	23.5	39.8
6.4	7.3	55.8	12.6	17.4	8.9	10.5	17.2	33.8
8.7	10.4	33.2	19.9	9.3	17.2	11.5	15.4	25.7
10.9	14.7	29.7	33.6	18.5	23.1	15.1	12.2	24.9
19.5	15.6	194.5	15.2	69.9	39.5	14.5	12.7	62.2
12.1	15.1	215.6	26.2	66.8	50.1	22.6	1.2	43.4
10.1	19.1	36.3	35.1	46.7	36.5	15.2	4.6	22.4
11.5	13.7	53.2	25.1	17.5	25.0	14.7	13.3	30.3
8.4	9.6	54.6	16.9	17.4	13.7	12.1	15.7	32.4

length of residence, does it represent an error in the method of measuring intelligence, or, looked at from another angle, does it show the gradually decreasing intelligence of the more recent immigrants examined in the army?

The hypothesis of growth of intelligence with increasing length of residence may be identified with the hypothesis of an error in the method of measuring intelligence, for we must assume that we are measuring *native or inborn intelligence*, and any increase in our test score due to any other factor may be regarded as an error. It is therefore necessary to examine two hypotheses, viz., (1) a defect in the measuring scale, and (2) a change in the character of the immigrants examined, in order to decide which is correct, or, in case both factors are operative, to estimate quantitatively the magnitude of one of the factors, so that allowance may be made for that factor and the weight of the other factor thus determined.

The most probable source of error in our measure of intelligence is that arising from the different types of examination. Examination alpha involves the use of English, and the ability to use English is a function of intelligence and education in its broadest sense. Examination beta involves no English, and the tests can not be considered as educational measures in any sense. The individual examinations were adapted to the linguistic ability of the person examined. We would therefore expect to find an error in two places only, first, in the selection of men for alpha and beta, and second in the relationship between alpha and beta as expressed on the combined scale.

If all members of our five years of residence groups had been given alpha, beta and individual examinations in equal proportions, then all would have been treated alike, and the relationship shown would stand without any possibility of error. But this is not the case. The actual per-

centage of men in each residence group taking each type of examination was as follows:

LENGTH OF RESIDENCE	ALPHA	BETA	STANFORD-BINET	PERFORMANCE SCALE
0 to 5 years....	19%	68%	2%	11%
6 to 10 years....	26%	65%	2%	7%
11 to 15 years....	41%	54%	1%	4%
16 to 20 years....	66%	32%	1%	1%
Over 20 years.....	73%	26%	½%	½%

Recognizing a variation in the type of examination given, our problem becomes that of determining whether or not any injustice has been done by converting results from these different types of examination into the combined scale. If the language and educational factors account for the rise in the average score on the combined scale with increasing years of residence, then we should expect that the contribution made to the combined scale score through examination beta would remain constant, and the contribution from examination alpha would increase very rapidly. On the other hand, if the rise is independent of the language and educational factor and due to the greater native intelligence of the older groups, we should expect the contribution from both types of examination to remain the same. In other words, if the combined scale is accurate, the disparity between proportions taking different types of examinations would make no difference in the final results. At the same time we would not expect to find the average scores on the combined scale made by way of alpha and beta to be the same, for beta was given not only to those who were illiterate, but also to the dull and stupid who failed to make a good score on alpha.

Computing, then, the average score on the combined scale made by way of examination alpha, beta, and the Stanford-Binet (which includes the results from the performance

scale) by each of the five years of residence groups, we find the following:

YEARS OF RESIDENCE GROUPS	SCORES DERIVED FROM ALPHA	SCORES DERIVED FROM BETA	SCORES DERIVED FROM THE STANFORD-BINET
0 to 5 years. . . .	13.46 (S.D. 2.47)	11.11 (S.D. 2.73)	9.99 (S.D. 2.21)
6 to 10 years. . .	13.57 (S.D. 2.45)	11.29 (S.D. 2.65)	9.86 (S.D. 2.22)
11 to 15 years. . .	13.91 (S.D. 2.25)	11.62 (S.D. 2.50)	10.19 (S.D. 2.09)
16 to 20 years. . .	14.31 (S.D. 2.19)	12.22 (S.D. 2.68)	. . . . .
Over 20 years. . .	14.56 (S.D. 2.32)	11.93 (S.D. 2.70)	. . . . .

There is a steady progress in the scores in examination alpha from "0 to 5 yrs." up to "over 20 yrs.," the total gain being 1.12 points on the combined scale. There is an equal amount of progress in the scores from examination beta, a gain of 1.11 points on the combined scale being made in a shorter period of time, i. e. from "0 to 5 yrs." to "16 to 20 yrs." If the increase in the average score on the combined scale from 11.41 to 13.82 were due to the language and educational factor, then the gain should come from alpha and not from beta, for alpha involves language and (indirectly) education, and beta does not. We actually find that the gain from each type of examination is about the same. This indicates, then, that the five years of residence groups are groups with real differences in native intelligence, and not groups laboring under more or less of a linguistic and educational handicap.

There remains but one hypothesis that might establish the fact that the increase in the score on the combined scale with increasing length of residence was due to an error in the measuring scale, and that is the hypothesis that the combined scale was constructed in such a fashion that it penalized individuals born in non-English speaking countries. It will be remembered that the combined scale was constructed from Group X, a special experimental group to which were given all three types of examination. Group

X was composed of 1047 individuals all of whom were born in English speaking countries.

Fortunately we are able to test the reliability of the combined scale under the most severe conditions. On page 654 of Memoir XV, Table 166 shows the scores on both alpha and beta of all individuals in Groups I, II and III of the principal sample who had been given both alpha and beta. This group includes 4893 cases. It is obvious that we may figure these 4893 cases as either alpha cases or beta cases and convert them into the combined scale either by Table 159 (the alpha conversion table) or Table 162 (the beta conversion table). This the army writers have done, and the results are given in Table 167 on page 655.

It is found that when we treat the 4893 cases as alpha cases the average score on the combined scale is 10.775 (S. D. 1.64). When we treat the same 4893 cases as beta cases, the average score on the combined scale is 12.158 (S. D. 2.63). The actual difference between the two averages is 1.383 (± 0.0298). In commenting on this result, the army writers state:

"At first glance these results seem rather startling, for one might suppose that going from alpha (for a given number of cases) to the combined scale ought to yield the same results as going from beta to combined scale. The facts are quite the contrary. However, this difference in no wise discredits the method. It must be remembered that in a group of this sort there is a large percentage of illiterates; thus the group no doubt includes a considerable proportion of the cases who made unsatisfactory scores in alpha and were recalled to beta not because of stupidity but because of language difficulty. When they reached beta, they were able to make scores more consistent with their ability. It is precisely this element of the group that causes the difference in the two means on the combined scale. The same

fact explains the wide differences in the standard deviations. The standard deviation of the combined scale distribution when reached by way of beta is larger than by way of alpha. Here the difference is no doubt due to the fact that in alpha both the stupid and the non-English speaking piled up in the lower class intervals, while in beta the stupid remained in the lower ranges and the more intelligent went higher, thus increasing the standard deviation." (p. 655.)

It is also possible to study the effect of using different conversion tables in the case of Group X, the special experimental group which was composed of individuals who were born in English speaking countries. The army writers report on page 645 the analysis of Group X in this manner. The following averages and standard deviations are reported for the different methods of treating the data:

	AVERAGE	S. D.
(1) Treating all 1047 cases as measured by alpha only...	13.82716	3.03940
(2) Treating all cases scoring less than 25 points in alpha as beta cases and the remainder as alpha cases.....	13.88606	3.03776
(3) Treating all cases scoring less than 50 points in alpha as beta cases and the remainder as alpha cases.....	13.94350	3.20690
(4) Treating all cases scoring less than 75 points in alpha as beta cases and the remainder as alpha cases.....	13.96981	3.30997

The approximate agreement of the averages derived by the four different methods is considered by the army writers to be a proof of the validity of the transformation tables which they publish for converting alpha, beta, and Stanford-Binet distributions into combined scale distributions.

We thus have two extreme instances of the results of

treating the same groups in different ways. If we take all cases in the principal sample to whom were given both alpha and beta, we find a difference of 1.383 (±0.0298) in the combined scale score of the group when figured first as alpha cases and then as beta cases. On the other hand, since the combined scale was empirically derived from the 1047 cases in Group X, it makes very little difference how we treat the results of that group. We must now determine whether the distributions of alpha scores of the foreign born groups most closely resemble the distribution of alpha scores in Group X, or the distribution of the alpha scores of the 4893 cases who had both alpha and beta. This question is very easily answered by turning to Table 7, which shows the alpha score distribution of the following groups:

(1) The 4893 cases who had both alpha and beta.
(2) The 1047 cases in the special experimental group.
(3) The 679 cases of foreign born who had been in this country from 0 to 5 years.
(4) The 1098 cases of foreign born who had been in this country from 6 to 10 years.
(5) The 1777 cases of foreign born who had been in this country from 0 to 10 years, a group obtained by combining (3) and (4).

All five distributions have been made comparable by reducing them to the proportion in each 1000 scoring at each class interval of the alpha examination. Figure 34 shows graphically the relationships between the 4893 alpha and beta cases, the 1047 cases in the special experimental group and the 1777 foreign born who had been in this country for 10 years or less. In Figure 34 the horizontal line indicates alpha scores and the vertical line shows the proportion in each 1000.

A hasty survey of Table 7 shows that our group of 4893 cases is a very specially selected group. 97.3% of the alpha

Table No. 7

Distribution of alpha scores of (1) all cases in Groups I, II, and III given both alpha and beta, (2) special experimental group, (3) foreign born in U. S. 0 to 5 years, (4) foreign born in U. S. 6 to 10 years, (5) foreign born in U. S. 0 to 10 years. All distributions reduced to common denominator of number per 1000 at each class interval. Actual distributions may be found on pages 621, 654 and 701 of Memoir XV.

ALPHA CLASS INTERVALS	4893 CASES TAKING BOTH ALPHA AND BETA	1047 CASES IN SPECIAL EXPERIMENTAL GROUP	679 CASES IN U. S. 0 TO 5 YEARS	1098 CASES IN U. S. 6 TO 10 YEARS	1777 CASES IN U. S. 0 TO 10 YEARS
185–189	....	3	....	....	....
180–184	....	1	2	1	1
175–179	....	1	....	....	....
170–174	....	2	4	1	2
165–169	....	7	2	1	1
160–164	....	4	....	3	2
155–159	....	5	2	2	2
150–154	....	6	....	4	2
145–149	....	7	6	2	3
140–144	....	9	12	5	7
135–139	1	12	6	8	7
130–134	....	9	3	7	6
125–129	....	20	7	7	7
120–124	1	13	3	6	5
115–119	....	19	9	11	10
110–114	1	12	19	10	14
105–109	....	21	9	14	12
100–104	1	19	15	17	17
95–99	1	39	25	16	19
90–94	1	24	25	27	26
85–89	1	30	22	37	31
80–84	2	41	27	40	34
75–79	1	25	32	31	31
70–74	2	39	53	29	39
65–69	1	38	29	46	40
60–64	8	44	47	51	50
55–59	7	38	52	48	50
50–54	6	35	66	55	59
45–49	8	40	72	56	62
40–44	12	44	68	75	73
35–39	19	53	81	75	78
30–34	23	53	49	61	56
25–29	59	48	57	68	64
20–24	104	53	44	62	55
15–19	149	31	29	27	28
10–14	286	42	22	29	26
5–9	207	46	29	21	24
0–4	104	67	72	47	57

scores of the 4893 cases are below a total alpha score of 50. On the other hand, the distributions of alpha scores in our groups of foreign born very closely resemble the distribution of alpha scores of Group X, the special experimental group on which the combined scale was based. If we compute the percentage of cases in each residence group graded on alpha falling below a total alpha score of 50, we find the following:

YEARS OF RESIDENCE GROUP	PER CENT. GRADED BY ALPHA AND FALLING BELOW A TOTAL ALPHA SCORE OF 50
0 to 5 years	52.4%
6 to 10 years	52.2%
11 to 15 years	48.7%
16 to 20 years	39.1%
Over 20 years	36.0%

Group X shows 47.7% of the 1047 cases falling below a score of 50 on alpha. This group is approximately the same as the group of foreign born we are studying. On the other hand, the group of 4893 cases shows 97.3% of the alpha scores below 50, and a very marked piling up of cases below 25 on alpha. The relationships between the distributions of scores in the three groups are very clearly shown in Figure 34.

It is clear that the combined scale would penalize the foreign born only if all individuals who took alpha and beta had been scored as alpha cases. As a matter of fact, the opposite is true, for every individual who took both alpha and beta was scored as a beta case, and the alpha score was disregarded. On this account our results on the foreign born are not subject to the distortion shown by treating 4893 alpha and beta cases as alpha cases.

The 4893 cases treated as beta cases, not as alpha cases, are in our group of 93,955 cases representing the white

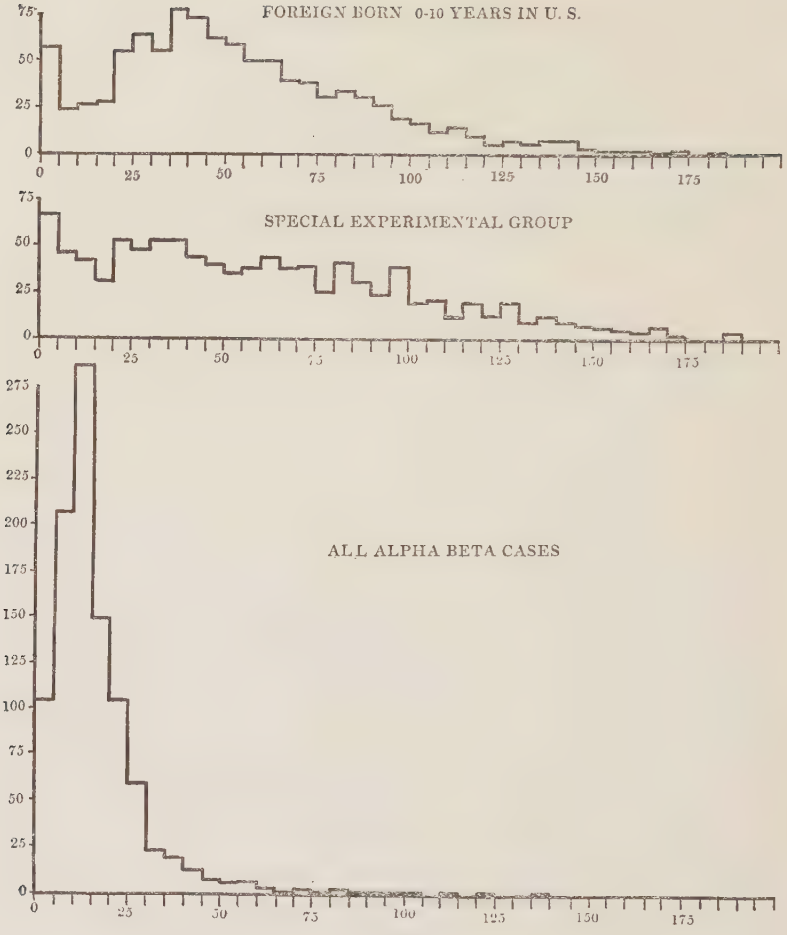
FOREIGN BORN 0-10 YEARS IN U. S.
75
50
25
0
0
25
50
75
100
125
150
175
SPECIAL EXPERIMENTAL GROUP
75
50
25
0
0
25
50
75
100
125
150
175
275
250
225
200
175
150
125
100
75
50
25
0
ALL ALPHA BETA CASES
0
25
50
75
100
125
150
175

Figure 34. Distributions of alpha scores of three groups: (A) foreign born individuals in this country from 0 to 10 years, (B) group X, the special experimental group, (C) all cases who took both alpha and beta. The horizontal scale shows alpha scores, and the vertical scale proportions in each thousand. The distributions are drawn from the figures given in Table 7. To prove that the combined scale does not penalize the foreign born, it is only necessary to show that the distribution of alpha scores for the foreign born is approximately similar to that of group X, and unlike that of all alpha-beta cases. This similarity is apparent from our chart. The men given both alpha and beta would have been penalized if they had been scored through the combined scale as alpha cases. This did not happen, however, for the alpha scores of these men were disregarded, and only the beta record used.

draft, and the chances are that most of them are among the 12,492 members of this group who were born in foreign countries. By treating the group as alpha cases, we show that the maximum correction factor that it would be necessary to apply to a group graded wrongly by alpha would be 1.383 points on the combined scale. But inasmuch as all alpha and beta cases have been correctly graded by using the beta conversion tables, there remain only those cases graded by alpha alone who should have been given examination beta as well. It is of course impossible to estimate the number of cases of this sort. If we make the very extravagant assumption that 25% of the cases graded by alpha alone in the residence group "0 to 5 yrs." should have been graded by beta, and credit each of these cases with 1.383 points, the final correction that it would be necessary to apply to our average would be 0.064 points, a quantity which is negligible in view of the magnitude of the differences under consideration. Furthermore, we should not be justified in using a correction factor of any sort unless it could be shown that the distribution of the alpha scores of the foreign born groups was very unlike that of the distribution of alpha scores of the special experimental group. As a matter of fact, these distributions as shown in Figure 34 are very much alike.

We have therefore demonstrated the accuracy of the combined scale as a measure of the intelligence of the groups under consideration. We must therefore accept the conclusion that under the conditions of this experiment the differences shown in the average scores of the five years of residence groups indicate real differences in intelligence and not a defect in the measuring scale. Instead of considering that our curve (Figure 33) indicates a growth of intelligence with increasing length of residence, we are forced to take the reverse of the picture and accept the

hypothesis that the curve indicates a gradual deterioration in the class of immigrants examined in the army, who came to this country in each succeeding five year period since 1902.

SECTION V

ANALYSIS OF IMMIGRATION TO THE UNITED STATES

The fact that the average intelligence of the immigrants examined in the army who came to this country in each successive five year period since 1902 becomes progressively lower with each succeeding period indicates that an explanation of this phenomenon might be found in a change in the character of immigration. We must, therefore, turn to the statistics on immigration to see if any such change can be detected. Table 8 shows the percentage of the total immigration coming from various countries in the periods roughly corresponding to the five years of residence periods covered in the army statistics. The data on immigration were obtained from the *Statistical Abstract of the United States* for the years 1900, 1910, and 1920. Table 8 reads as follows:—in the years 1887 to 1897, the period roughly corresponding with our residence group "over 20 yrs.," 10.9% of our total immigration came from England, 2.7% from Scotland, 1% from Holland, etc. The relations shown in Table 8 are shown graphically in Figure 35. Each complete bar in Figure 35 represents 100%. The per cent. which each country has contributed to the total immigration of each period has been scaled off proportionately in each bar.

These figures show that the most abrupt change in the character of immigration came between the periods 1887–1897 and 1898–1902. These periods show a very marked decrease in the proportion of the immigration from England and Germany, and a substantial decrease in the proportion of immigration from Scotland, Sweden, and

TABLE NO. 8

Per cent. of total immigration coming from various countries during periods roughly corresponding to the five years of residence groups.

	1913–1917 0 TO 5 YEARS	1908–1912 6 TO 10 YEARS	1903–1907 11 TO 15 YEARS	1898–1902 16 TO 20 YEARS	1887–1897 OVER 20 YEARS
England	3.7	5.1	4.6	2.6	10.9
Scotland	1.0	1.8	1.4	0.5	2.7
Holland	0.6	0.8	0.5	0.4	1.0
Germany	2.5	3.5	4.0	4.8	18.7
Denmark	0.7	0.7	0.8	0.8	1.7
Canada	13.5	6.0	0.6	0.2	No record
Sweden	1.5	2.0	2.8	4.6	7.3
Norway	1.1	1.5	2.3	2.4	2.7
Belgium	0.5	0.5	0.5	0.3	0.6
Ireland	2.4	3.3	3.8	7.2	11.9
Austria	16.7	21.8	24.9	23.6	11.6
Turkey	2.6	3.4	1.9	0.7	No record
Greece	3.7	2.5	1.8	1.1	0.2
Russia	17.8	18.3	18.3	17.8	12.0
Italy	20.0	20.2	23.7	25.8	12.4
All others	11.7	8.6	8.1	7.2	6.3

Ireland. On the other hand, the proportion of immigrants coming from Austria, Russia, and Italy showed a marked increase at this time. In general the following relations held:

England showed a decided drop in the proportion of immigrants furnished between the period 1887–1897 and the period 1898–1902. There has been a slight increase since 1898, but the proportion is less than 5%, when formerly it was over 10%.

Scotland contributed 2.7% of our total immigration in the period 1887–1897, and since that time, never more than 2%.

Holland never contributed more than 1% of our total immigration in any period covered by these figures.

Germany contributed 18.7% of our immigration in the period 1887–1897, 4.8% in the period 1898–1902, and since that time the proportion has decreased with each succeeding period.

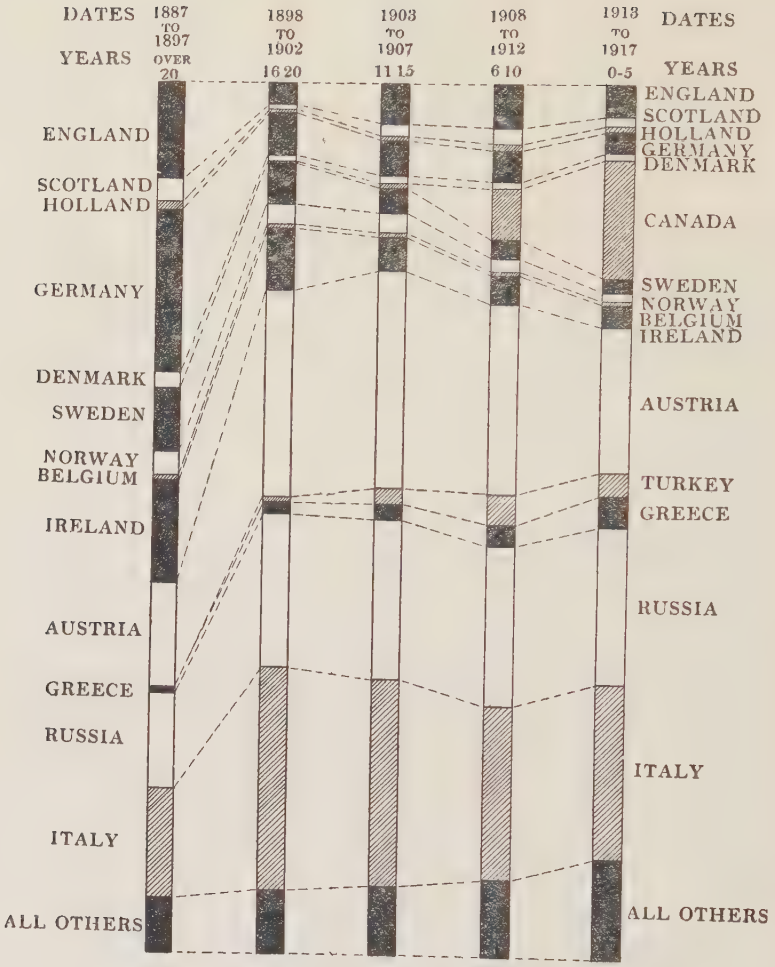
DATES
YEARS
1887 TO 1897 OVER 20
1898 TO 1902 16 20
1903 TO 1907 11 15
1908 TO 1912 6 10
1913 TO 1917 0-5
DATES
YEARS
ENGLAND
SCOTLAND
HOLLAND
GERMANY
DENMARK
SWEDEN
NORWAY
BELGIUM
IRELAND
AUSTRIA
GREECE
RUSSIA
ITALY
ALL OTHERS
CANADA
TURKEY

Figure 35. Analysis of immigration by countries. Each vertical bar represents 100%, and each subdivision represents the per cent. of the total immigration that each country furnished in the period under consideration. It is apparent that the most sudden change in the character of immigration came between the periods 1887-1897 and 1898-1902. On the other hand, the curve showing the relationship between years of residence and average intelligence shows no correspondingly large diffrence in the intelligence of immigrants who came here in these periods (Figure 33, Groups "16 to 20" and "over 20"). Figure 33 shows a large difference between the "16 to 20 yrs." and "11 to 15 yrs." groups, but the chart above shows no marked difference in the immigration coming to this country in the corresponding five-year periods, 1898-1902 and 1903-1907.

Denmark contributed 1.7% of our immigration in the period 1887–1897 and since that time less than 1% in each period.

The immigration from Canada was not reported prior to 1896. There was a very marked increase in the percentage of our total immigration coming from Canada in the period 1908–1912 and again in the period 1913–1917.

The proportion of immigration coming from Sweden has decreased steadily from 7.3% in the period 1887–1897 to 1.5% in the period 1913–1917. A similar decrease from 2.7% to 1.1% is shown by Norway during the same periods.

Belgium has contributed less than 1% to the total immigration in each period under consideration.

The immigration from Ireland has decreased in proportion from 11.9% to 2.4% in the five year periods shown.

All other countries show a gain in the proportion of immigrants which they supply. Austria supplied 11.6% in the period 1887–1897, this proportion jumping to 23.6% in the period 1898–1902 and remaining above 20% until the last period 1913–1917, a period which reflects the war conditions in Europe.

The proportion of immigration from Russia has increased from 12% prior to 1898 to 18% since that time. The proportion of immigration from Italy, which was about 12% prior to 1898, has never been below 20% since. Turkey and Greece show a small but increasing proportion in the successive periods covered.

Enough evidence has been cited to show that there has been some change in the character of our immigration during the periods covered in the army report. The gradual decline in the average intelligence of the more recent immigrants examined in the army might be due to these changes in the source of supply. If this hypothesis is correct, we should expect to find differences between the scores of

the representatives of each country making up our total group of the foreign born white draft.

SECTION VI

ANALYSIS OF THE FOREIGN BORN WHITE DRAFT BY COUNTRY OF BIRTH

The army reports show distributions of psychological test scores for all of the 12,492 foreign born cases classified according to the country of birth, although these figures for each country are not sub-divided again into years of residence groups. As a matter of fact, when we break up the 12,492 cases according to the country of birth, the figures from certain countries become very small; so that further sub-division would make the results valueless. We can, however, examine the figures which give us an intelligence measure of the foreign born men of our army and take them for what they are worth. Even though the number of cases is very small for certain nationalities, we previously saw that the reliability of the difference does not depend entirely on the number of cases, but on three factors—the size of the difference, the variability of each series of measurements, and the number of cases in each series.

The same tables from which we derived the 12,492 foreign born cases (Tables 207, 208, 209 and 210; pp. 692, 693 and 694) give the classification of test scores by country of birth. From these tabulations the combined scale score of each nativity group has been computed. The actual distributions of these combined scale scores are shown in Table 9. These distributions reduced to percentages are shown in Table 10.

The differences found, expressed in terms of the per cent. from each country who exceed the average native born American, are as follows:

England	67.3%	Belgium	35.3%
Scotland	58.8%	Austria	28.2%
Holland	58.1%	Ireland	26.2%
Germany	48.7%	Turkey	25.3%
Denmark	47.8%	Greece	21.3%
Canada	47.3%	Russia	18.9%
Sweden	41.7%	Italy	14.4%
Norway	37.3%	Poland	12.2%

The actual differences between the average scores on the combined scale of each country and every other country, together with the probable errors of the differences, and the ratios of the probable errors of the differences to the differences are shown in Tables 11 to 27 inclusive. In these tables, the convention has been followed of marking a difference "unreliable" if the actual difference was less than four times the probable error of the difference. The countries tabulated are arranged in the order of the increasing reliability of the differences found, above and below the average of the country with which they are compared.

The relations between the averages shown in Table 9 are shown graphically in Figure 36. For the convenience of those who use Stanford-Binet "mental ages," one side of the scale in Figure 36 has been drawn so as to read in "mental ages," and the other side to read in combined scale units—the units in which our averages and measures of variability have been calculated. The Stanford-Binet "mental age" scale was calculated from the regression equation given on page 654 of Memoir XV:

Mental age (in years) = 0.778 C.S. + 2.606.

Table No. 9

Analysis of foreign born white draft by country of birth. Distribution of scores on the combined scale, of men born in each country.

COMBINED SCALE INTERVALS	ENGLAND	SCOTLAND	HOLLAND	GERMANY	DENMARK	CANADA	SWEDEN	NORWAY
24.0–24.9	0.1	....	....	....	....	....	....	....
23.0–23.9	0.7	....	....	....	....	....	....	....
22.0–22.9	1.7	0.5	0.1	0.5	....	0.6	....	....
21.0–21.9	3.0	1.4	0.7	1.2	0.1	4.3	0.5	0.2
20.0–20.9	6.0	1.9	1.6	2.4	0.5	8.7	2.2	1.2
19.0–19.9	10.7	2.6	3.0	5.1	1.9	15.8	5.4	3.2
18.0–18.9	21.1	3.7	4.5	7.1	5.8	28.2	10.2	8.4
17.0–17.9	33.7	8.9	7.2	14.7	13.7	47.8	21.3	18.4
16.0–16.9	48.9	17.9	12.9	24.7	26.4	77.2	44.3	34.0
15.0–15.9	64.2	20.8	20.4	34.4	39.9	104.4	72.6	59.1
14.0–14.9	71.5	22.9	25.1	46.3	52.0	135.2	101.9	81.4
13.0–13.9	65.7	23.0	25.5	57.1	64.0	163.2	125.6	98.5
12.0–12.9	40.9	18.0	18.6	53.2	54.2	144.4	118.3	101.2
11.0–11.9	19.2	11.0	10.2	31.5	31.2	93.5	79.3	79.5
10.0–10.9	10.2	6.1	5.8	16.4	18.2	60.2	51.4	55.8
9.0–9.9	6.1	3.5	2.8	8.0	10.3	41.7	31.7	36.4
8.0–8.9	3.7	2.1	1.1	3.5	4.6	25.0	16.2	19.8
7.0–7.9	2.2	1.1	0.4	1.3	1.7	13.0	7.0	9.2
6.0–6.9	1.0	0.5	0.1	0.5	0.4	5.6	2.4	3.4
5.0–5.9	0.3	0.1	....	0.1	0.1	2.1	0.6	1.0
4.0–4.9	0.1	....	....	....	....	0.8	0.1	0.3
3.0–3.9	....	....	....	....	....	0.2	....	....
2.0–2.9	....	....	....	....	....	0.1	....	....
1.0–1.9	....	....	....	....	....	....	....	....
No. cases	411	146	140	308	325	972	691	611
Average	14.87	14.34	14.32	13.88	13.69	13.66	13.30	12.98
S. D.	2.57	2.63	2.39	2.43	2.23	2.67	2.38	2.47

BELGIUM	IRELAND	AUSTRIA	TURKEY	GREECE	RUSSIA	ITALY	POLAND
....	....	....	....	....	....	....	....
....	....	....	....	....	....	....	....
....	0.1	....	....	....	0.6	0.2	....
....	1.1	0.4	0.3	....	2.3	0.8	....
....	2.0	1.4	0.8	0.2	3.8	2.1	....
0.2	3.3	2.0	1.5	0.8	7.7	5.9	0.2
0.8	7.5	3.0	4.3	3.0	20.7	14.4	0.9
2.4	13.5	5.0	9.6	8.2	40.0	32.8	2.7
7.0	23.7	10.1	15.6	14.8	62.7	70.0	5.8
13.3	38.2	20.1	24.4	26.5	97.7	136.4	11.1
16.6	59.4	32.1	36.9	49.8	152.8	235.0	18.3
22.3	102.0	46.8	57.2	83.0	233.6	377.2	33.1
22.6	122.6	50.8	67.4	99.2	301.1	508.7	47.2
16.5	94.0	38.2	56.9	88.5	317.0	569.2	53.8
11.0	68.9	29.3	47.8	73.1	316.7	596.3	56.0
7.3	55.0	24.7	41.0	57.3	299.2	573.6	55.2
4.4	35.7	17.5	28.9	35.5	226.0	423.4	43.4
2.6	18.6	10.6	17.3	19.0	141.3	255.7	28.5
1.3	8.0	5.4	8.2	8.5	71.1	126.6	15.2
0.5	3.0	2.3	3.3	3.2	29.9	53.0	6.8
0.2	1.0	0.9	1.2	1.0	11.2	19.5	2.6
....	0.3	0.3	0.3	0.3	3.4	6.0	0.9
....	0.1	0.1	0.1	0.1	1.0	1.8	0.3
....	....	....	....	....	0.2	0.4	....
129	658	301	423	572	2340	4009	382
12.79	12.32	12.27	12.02	11.90	11.34	11.01	10.74
2.42	2.60	2.75	2.75	2.45	2.83	2.60	2.59

Table No. 10

Analysis of foreign born white draft by country of birth. Per cent. from each country scoring at each interval on the combined scale.

COMBINED SCALE INTERVALS	ENGLAND	SCOTLAND	HOLLAND	GERMANY	DENMARK	CANADA	SWEDEN	NORWAY
23.0–23.9	0.2	....	....	....	....	....	....	....
22.0–22.9	0.4	0.3	0.1	0.2	....	0.1	....	....
21.0–21.9	0.7	1.0	0.5	0.4	....	0.4	0.1	....
20.0–20.9	1.5	1.3	1.1	0.8	0.2	0.9	0.3	0.2
19.0–19.9	2.6	1.8	2.1	1.7	0.6	1.6	0.8	0.5
18.0–18.9	5.1	2.5	3.2	2.3	1.8	2.9	1.5	1.4
17.0–17.9	8.2	6.1	5.1	4.8	4.2	4.9	3.1	3.0
16.0–16.9	11.9	12.2	9.3	8.0	8.2	7.9	6.4	5.5
15.0–15.9	15.6	14.3	14.6	11.2	12.3	10.8	10.5	9.7
14.0–14.9	17.4	15.7	17.9	15.0	16.0	13.9	14.8	13.3
13.0–13.9	16.0	15.8	18.2	18.5	19.7	16.8	18.1	16.1
12.0–12.9	10.0	12.3	13.3	17.3	16.6	14.9	17.2	16.6
11.0–11.9	4.7	7.5	7.3	10.2	9.6	9.6	11.4	13.0
10.0–10.9	2.5	4.2	4.1	5.3	5.6	6.2	7.5	9.1
9.0–9.9	1.5	2.4	2.0	2.6	3.2	4.3	4.6	6.0
8.0–8.9	0.9	1.4	0.8	1.1	1.4	2.6	2.3	3.3
7.0–7.9	0.5	0.8	0.3	0.4	0.5	1.3	1.0	1.5
6.0–6.9	0.2	0.4	0.1	0.2	0.1	0.6	0.3	0.6
5.0–5.9	0.1	....	....	....	....	0.2	0.1	0.2
4.0–4.9	....	....	....	....	....	0.1	....	....
3.0–3.9	....	....	....	....	....	....	....	....
2.0–2.9	....	....	....	....	....	....	....	....

BELGIUM	IRELAND	AUSTRIA	TURKEY	GREECE	RUSSIA	ITALY	POLAND
....	....	....	....	....	....	....	....
....	....	....	....	....	....	....	....
....	0.2	0.1	0.1	....	0.1	....	....
....	0.3	0.5	0.2	....	0.2	....	....
0.2	0.5	0.7	0.4	0.1	0.4	0.1	0.1
0.6	1.1	1.0	1.0	0.5	0.9	0.4	0.2
1.9	2.1	1.7	2.3	1.4	1.7	0.8	0.7
5.4	3.6	3.4	3.7	2.6	2.6	1.7	1.5
10.3	5.8	6.6	5.8	4.7	4.2	3.4	2.9
12.9	9.0	10.6	8.7	8.7	6.5	5.8	4.8
17.3	15.5	15.6	13.5	14.5	10.1	9.4	8.7
17.5	18.7	16.9	15.9	17.4	12.9	12.7	12.3
12.8	14.3	12.7	13.4	15.4	13.5	14.2	14.1
8.5	10.5	9.7	11.3	12.8	13.5	14.9	14.6
5.6	8.3	8.2	9.7	10.0	12.8	14.3	14.4
3.4	5.4	5.8	6.8	6.2	9.7	10.5	11.4
2.0	2.8	3.5	4.1	3.3	6.0	6.4	7.5
1.0	1.2	1.8	1.9	1.5	3.0	3.1	4.0
0.4	0.5	0.8	0.8	0.6	1.3	1.7	1.8
0.2	0.2	0.3	0.3	0.2	0.5	0.5	0.7
....	....	0.1	0.1	0.1	0.1	0.1	0.2
....	....	....	....	....	....	....	0.1

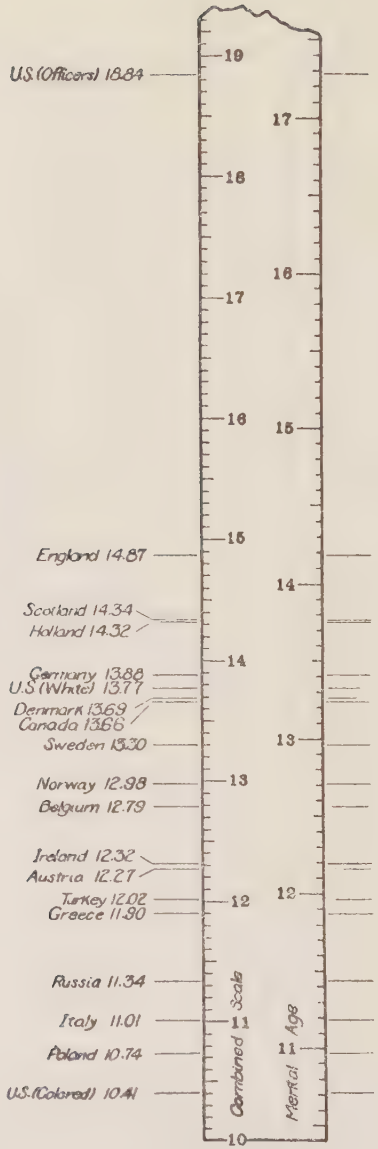
U.S.(Officers) 18.84
England 14.87
Scotland 14.34
Holland 14.32
Germany 13.88
U.S.(White) 13.77
Denmark 13.69
Canada 13.66
Sweden 13.30
Norway 12.98
Belgium 12.79
Ireland 12.32
Austria 12.27
Turkey 12.02
Greece 11.90
Russia 11.34
Italy 11.01
Poland 10.74
U.S.(Colored) 10.41
Combined Scale
Mental Age
19
18
17
16
15
14
13
12
11
10
17
16
15
14
13
12
11

Figure 36. The relative standing of the nativity groups according to their average intelligence. The averages of the nativity groups are taken from Table 9. The averages of the white officers and negro draft (from Table 1) and the native born white draft (from Table 2) are also shown. The left hand scale reads in units of the combined scale. The right hand scale reads in units of "mental age" representing what would be the approximately equivalent scores on the Stanford revision of the Binet-Simon scale. In interpreting the differences shown, it must be remembered that all the differences are not equally reliable, for the reliability of the measurements depends on the number of cases in each group and the variability of the group. The reliability of all the differences is shown in Tables 11 to 27.

Table No. 11

Differences between ENGLAND and other countries

Number of cases 411
Average score 14.87
Standard deviation 2.57

Scotland	−0.53 ±0.1695 (3.1)	Difference unreliable.
Holland	−0.55 ±0.1606 (3.4)	Difference unreliable.
Germany	−0.99 ±0.1264 (7.8)	
Denmark	−1.18 ±0.1192 (9.9)	
Canada	−1.21 ±0.1030 (11.7)	
Belgium	−2.08 ±0.1669 (12.5)	
United States	−1.10 ±0.0855 (12.9)	
Sweden	−1.57 ±0.1049 (14.9)	
Norway	−1.89 ±0.1087 (17.4)	
Austria	−2.60 ±0.1368 (19.0)	
Turkey	−2.85 ±0.1241 (22.9)	
Ireland	−2.55 ±0.1094 (23.3)	
Greece	−2.97 ±0.1097 (27.0)	
Poland	−4.13 ±0.1236 (33.4)	
Russia	−3.53 ±0.0940 (37.5)	
Italy	−3.86 ±0.0897 (43.0)	

Table No. 12

Differences between SCOTLAND and other countries

Number of cases 146
Average score 14.34
Standard deviation 2.63

England	+0.53 ±0.1695 (3.1)	Difference unreliable.
Holland	−0.02 ±0.1999 (0.1)	Difference unreliable.
Germany	−0.46 ±0.1736 (2.6)	Difference unreliable.
United States	−0.57 ±0.1466 (3.9)	Difference unreliable.
Denmark	−0.65 ±0.1685 (3.9)	Difference unreliable.
Canada	−0.68 ±0.1575 (4.3)	
Sweden	−1.04 ±0.1587 (6.5)	
Belgium	−1.55 ±0.2050 (7.6)	
Norway	−1.36 ±0.1612 (8.4)	
Austria	−2.07 ±0.1813 (11.4)	
Ireland	−2.02 ±0.1617 (12.5)	
Turkey	−2.32 ±0.1720 (13.5)	
Greece	−2.44 ±0.1619 (15.0)	
Russia	−3.00 ±0.1517 (19.8)	
Poland	−3.60 ±0.1716 (20.9)	
Italy	−3.33 ±0.1491 (22.3)	

Table No. 13

Differences between HOLLAND and other countries

Number of cases 140
Average score 14.32
Standard deviation 2.39

England	+0.55 ±0.1606 (3.4)	Difference unreliable.
Scotland	+0.02 ±0.1999 (0.1)	Difference unreliable.
Germany	−0.44 ±0.1649 (2.7)	Difference unreliable.
Denmark	−0.63 ±0.1595 (3.9)	Difference unreliable.
United States	−0.55 ±0.1362 (4.0)	Difference unreliable.
Canada	−0.66 ±0.1479 (4.4)	
Sweden	−1.02 ±0.1491 (6.8)	
Belgium	−1.53 ±0.1977 (7.7)	
Norway	−1.34 ±0.1518 (8.8)	
Austria	−2.05 ±0.1731 (11.8)	
Ireland	−2.00 ±0.1523 (13.1)	
Turkey	−2.30 ±0.1632 (14.1)	
Greece	−2.42 ±0.1525 (15.9)	
Russia	−2.98 ±0.1417 (21.0)	
Poland	−3.58 ±0.1628 (21.9)	
Italy	−3.31 ±0.1388 (23.9)	

Table No. 14

Differences between GERMANY and other countries

Number of cases 308
Average score 13.88
Standard deviation 2.43

England	+0.99 ±0.1264 (7.8)	
Holland	+0.44 ±0.1649 (2.7)	Difference unreliable.
Scotland	+0.46 ±0.1736 (2.6)	Difference unreliable.
United States	−0.11 ±0.0934 (1.1)	Difference unreliable.
Denmark	−0.19 ±0.1249 (1.5)	Difference unreliable.
Canada	−0.22 ±0.1097 (2.0)	Difference unreliable.
Sweden	−0.58 ±0.1114 (5.2)	
Belgium	−1.09 ±0.1711 (6.3)	
Norway	−0.90 ±0.1150 (7.8)	
Austria	−1.61 ±0.1418 (11.3)	
Ireland	−1.56 ±0.1157 (13.5)	
Turkey	−1.86 ±0.1297 (14.3)	
Greece	−1.98 ±0.1159 (17.1)	
Poland	−3.14 ±0.1291 (24.3)	
Russia	−2.54 ±0.1012 (25.1)	
Italy	−2.87 ±0.0972 (29.5)	

Table No. 15

Differences between the UNITED STATES and other countries. (Native born white draft used.)

Number of cases 81,465
Average score 13.77
Standard deviation 2.86

England	+1.10 ±0.0855 (12.9)	
Holland	+0.55 ±0.1362 (4.0)	Difference unreliable.
Scotland	+0.57 ±0.1466 (3.9)	Difference unreliable.
Germany	+0.11 ±0.0934 (1.1)	Difference unreliable.
Denmark	−0.08 ±0.0835 (1.0)	Difference unreliable.
Canada	−0.11 ±0.0582 (1.9)	Difference unreliable.
Belgium	−0.98 ±0.1437 (6.8)	
Sweden	−0.47 ±0.0614 (7.6)	
Norway	−0.79 ±0.0678 (11.6)	
Austria	−1.50 ±0.1071 (14.0)	
Turkey	−1.75 ±0.0904 (19.3)	
Ireland	−1.45 ±0.0688 (21.2)	
Greece	−1.87 ±0.0693 (26.9)	
Poland	−3.03 ±0.0896 (33.8)	
Russia	−2.43 ±0.0400 (60.7)	
Italy	−2.76 ±0.0285 (96.8)	

Table No. 16

Differences between DENMARK and other countries

Number of cases 325
Average score 13.69
Standard deviation 2.23

England	+1.18 ±0.1192 (9.9)	
Scotland	+0.65 ±0.1685 (3.9)	Difference unreliable.
Holland	+0.63 ±0.1595 (3.9)	Difference unreliable.
Germany	+0.19 ±0.1249 (1.5)	Difference unreliable.
United States	+0.08 ±0.0835 (1.0)	Difference unreliable.
Canada	−0.03 ±0.1013 (0.3)	Difference unreliable.
Sweden	−0.39 ±0.1032 (3.7)	Difference unreliable.
Belgium	−0.90 ±0.1659 (5.4)	
Norway	−0.71 ±0.1071 (6.6)	
Austria	−1.42 ±0.1355 (10.5)	
Ireland	−1.37 ±0.1079 (12.7)	
Turkey	−1.67 ±0.1228 (13.6)	
Greece	−1.79 ±0.1081 (16.5)	
Poland	−2.95 ±0.1221 (24.1)	
Russia	−2.35 ±0.0921 (25.5)	
Italy	−2.68 ±0.0877 (30.5)	

Table No. 17

Differences between CANADA and other countries

Number of cases 972
Average score 13.66
Standard deviation 2.67

England	+1.21 ±0.1030 (11.7)	
Holland	+0.66 ±0.1479 (4.4)	
Scotland	+0.68 ±0.1575 (4.3)	
Germany	+0.22 ±0.1097 (2.0)	Difference unreliable.
United States	+0.11 ±0.0582 (1.9)	Difference unreliable.
Denmark	+0.03 ±0.1013 (0.3)	Difference unreliable.
Sweden	−0.36 ±0.0840 (4.3)	
Belgium	−0.87 ±0.1547 (5.6)	
Norway	−0.68 ±0.0888 (7.6)	
Austria	−1.39 ±0.1215 (11.4)	
Ireland	−1.34 ±0.0896 (14.9)	
Turkey	−1.64 ±0.1071 (15.3)	
Greece	−1.76 ±0.0900 (19.6)	
Poland	−2.92 ±0.1064 (27.4)	
Russia	−2.32 ±0.0700 (33.1)	
Italy	−2.65 ±0.0641 (41.3)	

Table No. 18

Differences between SWEDEN and other countries

Number of cases 691
Average score 13.30
Standard deviation 2.38

England	+1.57	±0.1049 (14.9)	
United States	+0.47	±0.0614 (7.6)	
Holland	+1.02	±0.1491 (6.8)	
Scotland	+1.04	±0.1587 (6.5)	
Germany	+0.58	±0.1114 (5.2)	
Canada	+0.36	±0.0840 (4.3)	
Denmark	+0.39	±0.1032 (3.7)	Difference unreliable.
Belgium	−0.51	±0.1559 (3.3)	Difference unreliable.
Norway	−0.32	±0.0910 (3.5)	Difference unreliable.
Austria	−1.03	±0.1231 (8.4)	
Ireland	−0.98	±0.0918 (10.6)	
Turkey	−1.28	±0.1089 (11.8)	
Greece	−1.40	±0.0921 (15.2)	
Poland	−2.56	±0.1082 (23.6)	
Russia	−1.96	±0.0727 (26.9)	
Italy	−2.29	±0.0671 (34.1)	

Table No. 19

Differences between NORWAY and other countries

Number of cases 611
Average score 12.98
Standard deviation 2.47

England	+1.89 ±0.1087 (17.4)	
United States	+0.79 ±0.0678 (11.6)	
Holland	+1.34 ±0.1518 (8.8)	
Scotland	+1.36 ±0.1612 (8.4)	
Germany	+0.90 ±0.1150 (7.8)	
Canada	+0.68 ±0.0888 (7.6)	
Denmark	+0.71 ±0.1071 (6.6)	
Sweden	+0.32 ±0.0910 (3.5)	Difference unreliable.
Belgium	−0.19 ±0.1586 (1.2)	Difference unreliable.
Austria	−0.71 ±0.1264 (5.6)	
Ireland	−0.66 ±0.0961 (6.8)	
Turkey	−0.96 ±0.1126 (8.5)	
Greece	−1.08 ±0.0965 (11.2)	
Poland	−2.24 ±0.1119 (20.1)	
Russia	−1.64 ±0.0795 (20.6)	
Italy	−1.97 ±0.0729 (27.0)	

Table No. 20

Differences between BELGIUM and other countries

Number of cases 129
Average score 12.79
Standard deviation 2.42

England	+2.08 ±0.1669 (12.5)	
Holland	+1.53 ±0.1977 (7.7)	
Scotland	+1.55 ±0.2050 (7.6)	
United States	+0.98 ±0.1437 (6.8)	
Germany	+1.09 ±0.1711 (6.3)	
Canada	+0.87 ±0.1547 (5.6)	
Denmark	+0.90 ±0.1659 (5.4)	
Sweden	+0.51 ±0.1559 (3.3)	Difference unreliable.
Norway	+0.19 ±0.1586 (1.2)	Difference unreliable.
Ireland	−0.47 ±0.1590 (2.9)	Difference unreliable.
Austria	−0.52 ±0.1789 (2.9)	Difference unreliable.
Turkey	−0.77 ±0.1695 (4.5)	
Greece	−0.89 ±0.1592 (5.6)	
Russia	−1.45 ±0.1488 (9.8)	
Poland	−2.05 ±0.1691 (12.1)	
Italy	−1.78 ±0.1462 (12.2)	

Table No. 21

Differences between IRELAND and other countries

Number of cases 658
Average score 12.32
Standard deviation 2.60

England	+2.55 ±0.1094 (23.3)	
United States	+1.45 ±0.0688 (21.2)	
Canada	+1.34 ±0.0896 (14.9)	
Germany	+1.56 ±0.1157 (13.5)	
Holland	+2.00 ±0.1523 (13.1)	
Denmark	+1.37 ±0.1079 (12.7)	
Scotland	+2.02 ±0.1617 (12.5)	
Sweden	+0.98 ±0.0918 (10.6)	
Norway	+0.66 ±0.0961 (6.8)	
Belgium	+0.47 ±0.1590 (2.9)	Difference unreliable.
Austria	−0.05 ±0.1269 (0.4)	Difference unreliable.
Turkey	−0.30 ±0.1132 (2.7)	Difference unreliable.
Greece	−0.42 ±0.0972 (4.3)	
Russia	−0.98 ±0.0791 (12.4)	
Poland	−1.58 ±0.1126 (14.0)	
Italy	−1.31 ±0.0739 (17.7)	

Table No. 22

Differences between AUSTRIA and other countries

Number of cases 301
Average score 12.27
Standard deviation 2.75

England	+2.60	±0.1368 (19.0)	
United States	+1.50	±0.1071 (14.0)	
Holland	+2.05	±0.1731 (11.8)	
Canada	+1.39	±0.1215 (11.4)	
Scotland	+2.07	±0.1813 (11.4)	
Germany	+1.61	±0.1418 (11.3)	
Denmark	+1.42	±0.1355 (10.5)	
Sweden	+1.03	±0.1231 (8.4)	
Norway	+0.71	±0.1264 (5.6)	
Belgium	+0.52	±0.1789 (2.9)	Difference unreliable.
Ireland	+0.05	±0.1269 (0.4)	Difference unreliable.
Turkey	−0.25	±0.1398 (1.8)	Difference unreliable.
Greece	−0.37	±0.1272 (2.9)	Difference unreliable.
Russia	−0.93	±0.1139 (8.2)	
Poland	−1.53	±0.1393 (10.9)	
Italy	−1.26	±0.1104 (11.4)	

Table No. 23

Differences between TURKEY and other countries

Number of cases 423
Average score 12.02
Standard deviation 2.75

England	+2.85 ±0.1241 (22.9)	
United States	+1.75 ±0.0904 (19.3)	
Canada	+1.64 ±0.1071 (15.3)	
Germany	+1.86 ±0.1297 (14.3)	
Holland	+2.30 ±0.1632 (14.1)	
Denmark	+1.67 ±0.1228 (13.6)	
Scotland	+2.32 ±0.1720 (13.5)	
Sweden	+1.28 ±0.1089 (11.8)	
Norway	+0.96 ±0.1126 (8.5)	
Belgium	+0.77 ±0.1695 (4.5)	
Ireland	+0.30 ±0.1132 (2.7)	Difference unreliable.
Austria	+0.25 ±0.1398 (1.8)	Difference unreliable.
Greece	−0.12 ±0.1135 (1.0)	Difference unreliable.
Russia	−0.68 ±0.0984 (6.9)	
Poland	−1.28 ±0.1269 (10.1)	
Italy	−1.01 ±0.0944 (10.7)	

Table No. 24

Differences between GREECE and other countries

Number of cases 572
Average score 11.90
Standard deviation 2.45

England	+2.97 ±0.1097 (27.0)	
United States	+1.87 ±0.0693 (26.9)	
Canada	+1.76 ±0.0900 (19.6)	
Germany	+1.98 ±0.1159 (17.1)	
Denmark	+1.79 ±0.1081 (16.5)	
Holland	+2.42 ±0.1525 (15.9)	
Sweden	+1.40 ±0.0921 (15.2)	
Scotland	+2.44 ±0.1619 (15.0)	
Norway	+1.08 ±0.0965 (11.2)	
Belgium	+0.89 ±0.1592 (5.6)	
Ireland	+0.42 ±0.0972 (4.3)	
Austria	+0.37 ±0.1272 (2.9)	Difference unreliable.
Turkey	+0.12 ±0.1135 (1.0)	Difference unreliable.
Russia	−0.56 ±0.0795 (7.1)	
Poland	−1.16 ±0.1129 (10.3)	
Italy	−0.89 ±0.0744 (11.9)	

Table No. 25

Differences between RUSSIA and other countries

Number of cases 2340
Average score 11.34
Standard deviation 2.83

United States	+2.43	±0.0400 (60.7)
England	+3.53	±0.0940 (37.5)
Canada	+2.32	±0.0700 (33.1)
Sweden	+1.96	±0.0727 (26.9)
Denmark	+2.35	±0.0921 (25.5)
Germany	+2.54	±0.1012 (25.1)
Holland	+2.98	±0.1417 (21.0)
Norway	+1.64	±0.0795 (20.6)
Scotland	+3.00	±0.1517 (19.8)
Ireland	+0.98	±0.0791 (12.4)
Belgium	+1.45	±0.1488 (9.8)
Austria	+0.93	±0.1139 (8.2)
Greece	+0.56	±0.0795 (7.1)
Turkey	+0.68	±0.0984 (6.9)
Poland	−0.60	±0.0977 (6.1)
Italy	−0.33	±0.0483 (6.9)

Table No. 26

Differences between ITALY and other countries

Number of cases 4009
Average score 11.01
Standard deviation 2.60

United States	+2.76	±0.0285 (96.8)	
England	+3.86	±0.0897 (43.0)	
Canada	+2.65	±0.0641 (41.3)	
Sweden	+2.29	±0.0671 (34.1)	
Denmark	+2.68	±0.0877 (30.5)	
Germany	+2.87	±0.0972 (29.5)	
Norway	+1.97	±0.0729 (27.0)	
Holland	+3.31	±0.1388 (23.9)	
Scotland	+3.33	±0.1491 (22.3)	
Ireland	+1.31	±0.0739 (17.7)	
Belgium	+1.78	±0.1462 (12.2)	
Austria	+1.26	±0.1104 (11.4)	
Turkey	+1.01	±0.0944 (10.7)	
Greece	+0.89	±0.0744 (11.9)	
Russia	+0.33	±0.0483 (6.9)	
Poland	−0.27	±0.0936 (2.9)	Difference unreliable.

Table No. 27

Differences between POLAND and other countries

Number of cases 382
Average score 10.74
Standard deviation 2.59

United States	+3.03	±0.0896 (33.8)	
England	+4.13	±0.1236 (33.4)	
Canada	+2.92	±0.1064 (27.4)	
Germany	+3.14	±0.1291 (24.3)	
Denmark	+2.95	±0.1221 (24.1)	
Sweden	+2.56	±0.1082 (23.6)	
Holland	+3.58	±0.1628 (21.9)	
Scotland	+3.60	±0.1716 (20.9)	
Norway	+2.24	±0.1119 (20.1)	
Ireland	+1.58	±0.1126 (14.0)	
Belgium	+2.05	±0.1691 (12.1)	
Austria	+1.53	±0.1393 (10.9)	
Greece	+1.16	±0.1129 (10.3)	
Turkey	+1.28	±0.1269 (10.1)	
Russia	+0.60	±0.0977 (6.1)	
Italy	+0.27	±0.0936 (2.9)	Difference unreliable.

Tables 11 to 27 give the most accurate interpretation of the differences found between the various nativity groups that it is possible to derive from the army data. It is desirable, however, to attempt to estimate the meaning of these differences in terms of standards which have some popular significance. For this reason, the combined scale distributions in this study have been converted into estimates of the per cent. of A and B men in each group, and the per cent. of D, D− and E men.

It should be remembered that the army letter ratings are arbitrary ratings and have no real significance aside from the tests from which they were derived. The army rating "A" represents a certain score on the tests that should have been reached by 4% or 5% of the whole army group. In the same way the rating "B" was fixed so as to include the next 8% or 10%. It is of course absurd to deplore the fact that only 4% or 5% of the army were A men, when A was fixed so that only 4% or 5% could receive that rating. At the other end of the scale, the ratings D and D− were fixed so that they would include approximately 20% of the total group, and the E rating was reserved for those recommended for development battalions, special service organizations, rejection or discharge. The estimates made at the time the examinations were being standardized proved to be about right. The A and B groups which should have included 12% to 15% of the draft actually included 12%, and the D, D− and E groups, which should have included 20% to 25%, actually included 24%.

Table 28 gives the per cent. of cases in each nativity group who would be classified A or B according to the criterion of the upper 12% of the total unselected white draft. Table 29 gives the per cent. who would be classified as D, D− and E according to the criterion of 24% of the unselected group. The relations between the various nativ-

ity groups given in Tables 28 and 29 are shown graphically in Figure 37.

Another criterion that probably represents intellectual ability of a high order is the per cent. at or above the average of the white officers. The classification of the nativity groups according to this criterion is given in Table 30. At the other end of the scale, a criterion of inferiority that has a certain social significance is the per cent. at or below the average of the negro draft. Table 31 shows the nativity groups classified according to this criterion. The results given in Tables 30 and 31 are shown graphically in Figures 38 and 39.

It is not possible to determine accurately the percentage

TABLE NO. 28

Per cent. of each nativity group in the A and B groups

England	19.0
Native Born White Draft	13.2
Scotland	13.1
Holland	12.4
Total White Draft	12.0
Canada	11.1
Germany	10.1
Denmark	7.0
Sweden	5.9
Norway	5.3
Foreign Born White Draft	4.6
Ireland	4.3
Austria	4.1
Turkey	4.0
Russia	3.3
Belgium	2.9
Greece	2.2
Italy	1.5
Colored Draft	1.4
Poland	1.1

of feeble-mindedness in each group. The selection by the draft boards probably excluded all idiots and many imbeciles. The diagnosis of the border-line cases of feeble-mindedness is, in the last analysis, a social diagnosis, and can not be based on intelligence tests alone. It has been found, however, that a "mental age" of eight indicates an order of intelligence so low that the individual has difficulty in adjusting himself to his environment adequately. Table 32 gives a conservative estimate of the per cent. in each nativity group below an approximate "mental age" of eight. The percentages in Table 32 are shown graphically in Figure 40.

TABLE NO. 29

Per cent. of each nativity group in the D, D−, and E groups

England	8.8
Holland	12.0
Scotland	13.5
Germany	16.2
Denmark	17.0
Native Born White Draft	21.0
Canada	21.6
Sweden	23.2
Total White Draft	24.1
Norway	29.0
Belgium	29.3
Ireland	38.1
Austria	38.4
Turkey	43.6
Greece	44.6
Foreign Born White Draft	44.6
Russia	55.7
Italy	60.5
Poland	63.8
Colored Draft	67.5

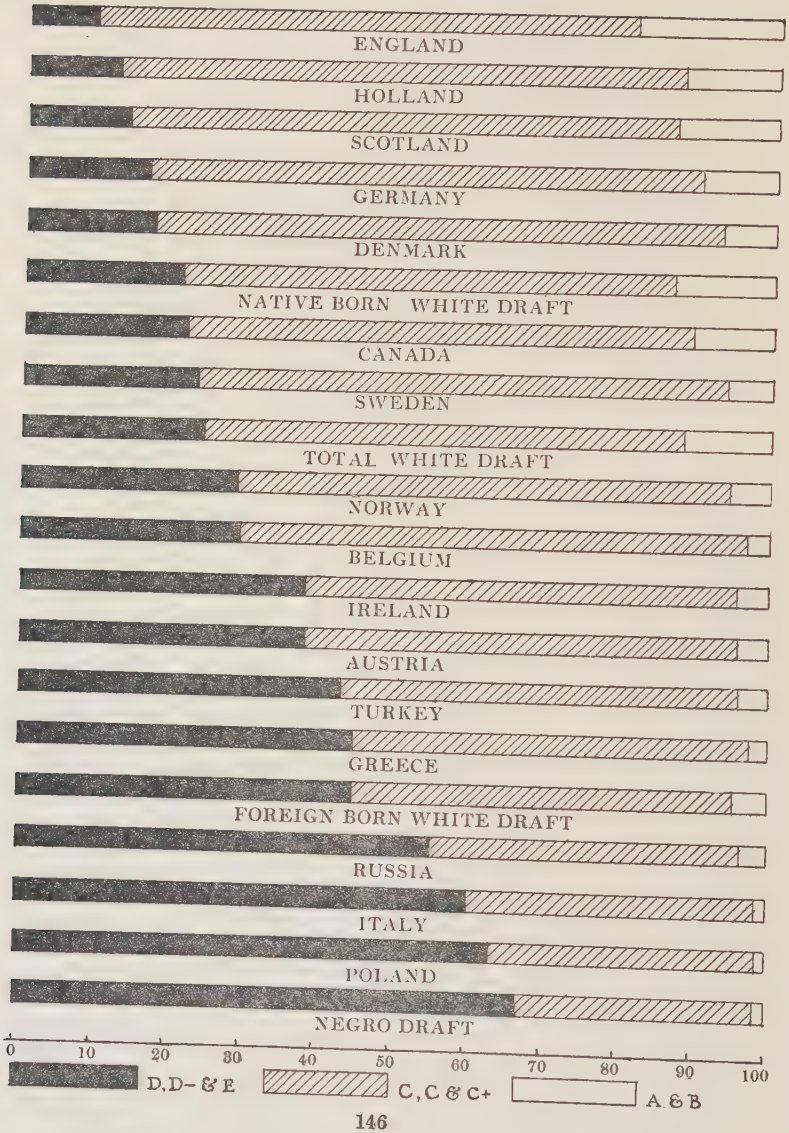
ENGLAND
HOLLAND
SCOTLAND
GERMANY
DENMARK
NATIVE BORN WHITE DRAFT
CANADA
SWEDEN
TOTAL WHITE DRAFT
NORWAY
BELGIUM
IRELAND
AUSTRIA
TURKEY
GREECE
FOREIGN BORN WHITE DRAFT
RUSSIA
ITALY
POLAND
NEGRO DRAFT
0
10
20
30
40
50
60
70
80
90
100
D, D– & E
C, C & C+
A & B

Figure 37. The relative standing of the various nativity groups in the proportions of A and B men, and D, D—, and E men. In interpreting this chart, it should be remembered that A and B, and D, D—, and E do not represent absolute intelligence standards, but rather standards arbitrarily fixed. In this case the standards were fixed by the 93,955 individuals making up the sample of the total white draft, A and B representing scores obtained by the upper 12% of this group, while D, D—, and E represent scores obtained by the lower 24%. The comparison is relative and not altogether reliable, for it fails to take into consideration the average, the number of cases, and the variability. Tables 28, 29, and this chart have been presented for the convenience of the reader, and to supplant Table 217, and Figure 19, on pages 697 and 698 of *Memoir XV,* which are not based on combined scale results.

Table No. 30

Per cent. of each nativity group at or above the average of the white officers

England	6.2
Scotland	4.8
Native Born White Draft	4.6
Holland	4.3
Total White Draft	4.1
Canada	3.5
Germany	3.4
Austria	1.5
Sweden	1.4
Foreign Born White Draft	1.3
Ireland	1.2
Denmark	1.1
Norway	1.0
Turkey	0.8
Russia	0.8
Belgium	0.3
Italy	0.3
Greece	0.3
Poland	0.1

Figure 38. The proportion of each nativity group obtaining intelligence scores at or above the average of the white officers (18.84). Reference to Figure 31 will show that this criterion indicates a relatively high order of intelligence. In comparing this Figure with Figures 37 and 39, it should be noted that each of the three figures has been drawn to a different scale. Our interpretation of these figures must be made with caution, for we are comparing extremes of the distribution curves without reference to the position of the average, the variability about the average, or the number of cases in the various groups.

0% 1% 2% 3% 4% 5% 6% 7%
ENGLAND
SCOTLAND
NATIVE BORN WHITE DRAFT
HOLLAND
TOTAL WHITE DRAFT
CANADA
GERMANY
AUSTRIA
SWEDEN
FOREIGN BORN WHITE DRAFT
IRELAND
DENMARK
NORWAY
TURKEY
RUSSIA
BELGIUM
ITALY
GREECE
POLAND
0% 1% 2% 3% 4% 5% 6% 7%

Table No. 31

Per cent. of each nativity group at or below the average of the negro draft

England	4.3
Holland	4.9
Germany	6.5
Scotland	6.8
Denmark	7.5
Native Born White Draft	7.6
Sweden	11.5
Canada	11.6
Total White Draft	13.7
Norway	15.2
Belgium	16.0
Ireland	22.8
Austria	24.5
Greece	27.1
Turkey	28.2
Foreign Born White Draft	29.5
Russia	39.0
Italy	42.3
Poland	46.0

Figure 39. The proportion of each nativity group at or below the average of the negro draft. Reference to Figure 31 will show that this criterion indicates a rather low order of intelligence. If 50% of any nativity group were at or below the negro average, the two distributions would be approximately identical. Russia shows 39% below the negro average, Italy 42.3%, and Poland 46%.

0%
10%
20%
30%
40%
50%
ENGLAND
HOLLAND
GERMANY
SCOTLAND
DENMARK
NATIVE BORN WHITE DRAFT
SWEDEN
CANADA
TOTAL WHITE DRAFT
NORWAY
BELGIUM
IRELAND
AUSTRIA
GREECE
TURKEY
FOREIGN BORN WHITE DRAFT
RUSSIA
ITALY
POLAND
0%
10%
20%
30%
40%
50%

Table No. 32

Per cent. of each nativity group below the approximate "mental age" of eight

Holland	0.1
Germany	0.2
Denmark	0.2
England	0.3
Scotland	0.4
Sweden	0.4
Norway	0.8
Canada	0.9
Native Born White Draft	1.1
Total White Draft	1.4
Belgium	1.6
Ireland	1.9
Greece	2.3
Austria	2.7
Turkey	3.1
Foreign Born White Draft	3.2
Russia	5.0
Italy	5.2
Poland	6.8
Colored Draft	10.0

Figure 40. Proportion of each nativity group testing below the approximate "mental age" of eight. This criterion indicates intelligence of a very low order. These individuals are probably capable of adjusting themselves only to the simplest form of environment, occupation and conditions of living. Few of them would be able to manage their affairs with ordinary prudence. Many of them should be in custodial institutions.

HOLLAND
GERMANY
DENMARK
ENGLAND
SCOTLAND
SWEDEN
NORWAY
CANADA
NATIVE BORN WHITE DRAFT
TOTAL WHITE DRAFT
BELGIUM
IRELAND
GREECE
AUSTRIA
TURKEY
FOREIGN BORN WHITE DRAFT
RUSSIA
ITALY
POLAND
NEGRO DRAFT
0%
1%
2%
3%
4%
5%
6%
7%
8%
9%
10%
11%

SECTION VII

RELIABILITY OF THE RESULTS

The results of the army psychological examination figured by means of the combined scale give us the best available measures of the intelligence of the various groups examined. Do these results apply to the army as a whole? The logic underlying the answer to this question is the same as that underlying the judgment of the whole by one of its parts. The tea taster *samples* the tea to be graded. He does not need to brew a whole bale of tea to find its worth. In this experiment we have sampled the entire army by taking 15,543 white officers, 93,955 white recruits and 23,596 negro recruits. Our group of white recruits was subdivided into 81,465 native born and 12,492 foreign born.

No one would hesitate to accept the results of the 81,465 native born as typical of the army as a whole. If we continued sampling indefinitely, our results would increase in reliability only as the square root of the number of cases, and 81,465 cases constitute a sample that is a luxury from the point of view of size. In the same way, no one could seriously question the reliability of our sampling of 15,543 officers, 23,596 negroes and 12,492 foreign born as typical of the remainder of the officers, negroes, and foreign born whites in the army.

The results from the 81,465 cases in the native born white draft may be taken as typical of white males between the ages of 21 and 31 and above the idiot or imbecile grade. In making our comparisons between other groups, we know that the Selective Service Act called all men to the colors impartially. The same regulations drew the Italians, the negroes, the native whites, the Polish, and all other

groups into the army. The method of sampling all the psychological records again drew these cases impartially. If our theory of sampling is correct, then we may accept the army results as very approximately typical of the male population as a whole.

For instance, our figures in Table 3 show characteristic differences in the average score on the army tests of foreign born individuals in this country from 0 to 5 yrs., and those in this country from 6 to 10 yrs., etc. The same factors which determined the sampling of the 3,576 cases in this country 0 to 5 yrs. determined the sampling of the 4,287 cases in this country 6 to 10 yrs. As long as the principles of sampling are the same, we may take our small sample as typical of the group as a whole.

The results of the psychological tests of foreign born individuals classified according to length of residence, taken as typical of our foreign born population as a whole, indicate definitely that the average intelligence of succeeding waves of immigration has become progressively lower. Immigrants coming to this country between 1913 and 1917 have a lower average intelligence than those coming to this country in the years 1908 to 1912. The group coming to this country in the years 1903 to 1907 had a higher average intelligence that the 1908 to 1912 group, and a lower average intelligence than immigrants coming to this country in the years 1898 to 1902.

In drawing these conclusions we are taking the groups examined in the army as typical of the corresponding groups in the entire population. During the years 1913 to 1917, about 3 1/3 millons of immigrants came to this country. We are actually using 3,576 cases or about 0.1% as typical of the whole group. It may very properly be objected that this is too small a sample on which to base definite conclusions. We must therefore state our conclu-

sions less dogmatically, and with the proviso that if the groups examined in the army are typical of the immigration coming to this country in the same periods, then we know that our more recent periods of immigration give us an average intelligence which becomes progressively lower and lower. This tentative conclusion will be modified by any evidence which tends to support the hypothesis made.

The same kind of argument from the sample to the group holds in our interpretation of the differences in the average intelligence scores of groups in our army born in different countries. For instance, in the period under consideration from 1887 to 1917 there have been about 3 7/8 millons of Italians, and over 3 million Russians who have come to this country. We are actually using 4,009 Italians and 2,340 Russians as typical samples of these groups. Of course no one would maintain that these 4,009 Italians are typical of the population of Italy. There are so many variable factors determining immigration that the immigrants can not themselves be taken as representative of the country as a whole. The question at issue is that of accepting 4,009 Italians as typical of the 3 7/8 millions who have come to this country since 1887. The chief claim to reliability of our sample from each country is the fact that the sample was drawn at random from the army group, and the fact that the Selective Service Act drew the men from each country impartially.

SECTION VIII

THE RACE HYPOTHESIS

The results of the examination of the nativity groups suggest immediately that the race factor may underlie the large differences found. If we do find the common factor of race underlying the differences between the various nativity groups, it will give our results much greater reliability, for the chance factors of sampling particularly inferior or superior groups in the small nativity samples would disappear in combination. Our figures are based on country of birth and no statistics are available for race. The race hypothesis must therefore be examined indirectly.

Writers on immigration, for the most part, divide the countries of Europe into two groups (1) Northern and Western, and (2) Eastern and Southern, and usually assume that the immigration from Northern and Western Europe has been mostly Nordic. This traditional method is open to two very serious objections. In the first place, the classification fails to differentiate the Alpine and Mediterranean race groups. In the second place the assumption that the immigration from Northern and Western Europe was mostly of a pure Nordic type is unwarranted, for this classification includes Germany and Ireland, two countries that have contributed very largely to our immigration in the past. The following figures show the size of the Irish and German immigration:

DECADE	TOTAL IMMIGRATION	PER CENT. FROM IRELAND	PER CENT. FROM GERMANY	PER CENT. FROM IRELAND AND GERMANY
1820–1830	143,439	35%	5%	40%
1831–1840	599,125	35%	25%	60%
1841–1850	1,171,251	46%	25%	71%
1851–1860	2,598,214	35%	37%	72%
1861–1870	2,314,824	19%	34%	53%
1871–1880	2,812,191	15%	26%	41%
1881–1890	5,246,613	12%	28%	40%
1891–1900	3,844,420	11%	14%	25%
1901–1910	8,795,386	4%	4%	8%
1911–1920	5,735,811	2½%	2½%	5%

These figures show clearly the fallacy of assuming that the immigration from Northern and Western Europe has been predominately Nordic, for Ireland is largely Mediterranean and Germany largely Alpine.

If we wish to obtain even approximate estimates of the contributions of each of the three European races to our importations, it is necessary to abandon the Northern and Western, and Eastern and Southern classification and try another method. If it were possible to make even approximate estimates of the percentage of Nordic, Alpine and Mediterranean blood in each of the European nations sending immigrants to this country, such approximate estimates would be very much superior to the present method.

In collaboration with students of this subject, I have constructed Table 33 which contains tentative estimates of the present blood constitution of the countries sending immigrants to this country. This table is, of course, only an approximation to the truth, and many persons will disagree with the estimates. For this reason, I am re-publishing in Table 34, Table 68, page 100, of the *Statistical Abstract for the United States* for 1920, which shows the arrivals of alien passengers and immigrants by nationalities and by

Table No. 33

Tentative estimates of the proportion of Nordic, Alpine and Mediterranean blood in each of the European countries.

	PER CENT. NORDIC	PER CENT. ALPINE	PER CENT. MEDITERRANEAN
Austria-Hungary	10	90	0
Belgium	60	40	0
Denmark	85	15	0
France	30	55	15
Germany	40	60	0
Greece	0	15	85
Italy	5	25	70
Netherlands	85	15	0
Norway	90	10	0
Sweden	100	0	0
Russia (including Poland)	5	95	0
Poland	10	90	0
Spain	10	5	85
Portugal	5	0	95
Roumania	0	100	0
Switzerland	35	65	0
Turkey (unclassified)	0	20	80
Turkey (in Europe) (including Serbia, Montenegro and Bulgaria)	0	60	40
Turkey (in Asia)	0	10	90
England	80	0	20
Ireland	30	0	70
Scotland	85	0	15
Wales	40	0	60
British North America	60	40	0

TABLE NO. 34

No. 68.—ARRIVALS OF ALIEN PASSENGERS AND IMMIGRANTS, 1820 TO 1920: BY NATIONALITIES AND BY DECADES.

[Sources: Records of the Bureau of Statistics prior to 1896; subsequently, reports of the Commissioner General of Immigration, Department of Labor. The figures represent "alien passengers" from Oct. 1, 1820, to Dec. 31, 1867; "immigrants" from Jan. 1, 1868, to date.

COUNTRY OF LAST PERMANENT RESIDENCE	OCT. 1, 1820, TO SEPT. 30, 1830	OCT. 1, 1830, TO DEC. 31, 1840	JAN. 1, 1841, TO DEC. 31, 1850	JAN. 1, 1851, TO DEC. 31, 1860.
Austria-Hungary	.....	.....	.....	.....
Belgium	27	22	5,074	4,738
Denmark	169	1,063	539	3,749
France	8,497	45,575	77,262	76,358
Germany	6,761	152,454	434,626	951,667
Greece[1]	.....	.....	.....	.....
Italy	408	2,253	1,870	9,231
Netherlands	1,078	1,412	8,251	10,789
Norway	91	1,201	13,903	20,931
Sweden				
Russia, including Russian Poland[2]	91	646	656	1,621
Spain[3]	2,622	2,954	2,759	10,353
Portugal[4]				
Rumania[1]	.....	.....	.....	.....
Switzerland	3,226	4,821	4,644	25,011
Turkey in Europe[5]	.....	.....	.....	.....
United Kingdom:				
England	22,167	73,143	263,332	385,643
Scotland	2,912	2,667	3,712	38,331
Ireland	50,724	207,381	780,719	914,119
Wales[6]	.....	.....	.....	.....
Total United Kingdom	75,803	283,191	1,047,763	1,338,093
Europe, not specified	43	96	155	116
Total Europe	98,816	495,688	1,597,502	2,452,657
British North America[7]	2,277	13,624	41,723	59,309
Mexico[7]	4,817	6,599	3,271	3,078
Central America	105	44	368	449
West Indies: Bermuda and Miquelon	3,834	12,301	13,528	10,660
South America	531	856	3,579	1,224
Total America[7]	11,564	33,424	62,469	74,720
Islands of the Atlantic	352	103	337	3,090
China	2	8	35	41,397
India[8]	.....	.....	.....	.....
Japan[8]	.....	.....	.....	.....
Turkey in Asia[8]	.....	.....	.....	.....
Other Asia	8	40	47	61
Total Asia	10	48	82	41,458
Total Oceania	2	9	29	158
Total Africa	16	52	55	210
All other countries	32,679	69,801	52,777	25,921
Grand total	143,439	599,125	1,713,251	2,598,214

[1]Included in "Europe, not specified," prior to 1891–1900. [2]Includes also Finland after 1872.
[3]Includes Canary and Balearic Islands after 1900.
[4]Figures include the Azores and Cape Verde Islands after 1879, they being classed with Portugal so far as that country is separately shown.

JAN. 1, 1861, TO JUNE 30, 1870	YEARS ENDED JUNE 30—				
	1871 TO 1880	1881 TO 1890	1891 TO 1900	1901 TO 1910	1911 TO 1920
7,800	72,969	353,719	597,047	2,145,266	896,342
6,784	7,221	20,177	20,062	41,635	33,746
17,094	31,771	88,132	52,670	65,285	41,983
35,984	72,206	50,464	36,006	73,379	61,897
787,468	718,182	1,452,970	543,922	341,498	143,945
.....	.....	.....	15,996	167,519	184,201
11,728	55,759	307,309	655,694	2,045,877	1,109,524
9,102	16,541	53,701	31,816	48,262	43,718
109,298	211,245	568,362	95,265	190,505	66,395
			230,679	249,534	95,074
4,536	52,254	265,088	593,703	1,597,306	921,957
8,493	9,893	6,535	6,723	27,935	68,611
			23,010	69,149	89,732
.....	.....	.....	14,559	53,008	13,311
23,286	28,293	81,988	33,149	34,922	23,091
.....	.....	.....	2,562	118,202	77,098
568,128	460,479	657,488	271,094	388,017	249,944
38,768	87,564	149,869	60,053	120,469	78,601
435,778	436,871	655,482	403,496	339,065	145,937
.....	.....	.....	11,186	17,464	13,107
1,042,674	984,914	1,462,839	745,829	865,015	487,589
210	656	10,318	4,370	1,719	18,350
2,064,407	2,261,904	4,721,602	3,703,061	8,136,016	4,376,564
153,871	383,269	392,802	2,631	179,226	742,185
2,191	5,362	1,913	746	49,642	219,004
96	210	462	1,183	8,112	17,159
9,043	13,957	29,042	35,040	107,548	123,424
1,396	928	2,304	3,059	17,280	41,899
166,597	403,726	426,523	42,659	361,808	1,143,671
3,446	10,056	15,798	.....	.....	.....
64,301	123,201	61,711	23,166	20,605	21,278
.....	.....	.....	26	4,713	2,082
.....	.....	.....	26,855	129,797	83,837
.....	.....	.....	8,398	77,393	79,389
308	622	6,669	28,370	11,059	5,973
64,609	123,823	68,380	86,815	243,567	192,559
221	10,913	12,574	8,793	12,973	13,427
312	229	437	1,343	7,368	8,443
15,232	1,540	1,299	1,749	33,654	1,147
2,314,824	2,812,191	5,246,613	3,844,420	8,795,386	5,735,811

[5]Includes Serbia, Bulgaria, and Montenegro prior to 1920; included in "Europe, not specified," prior to 1891–1900; also, after 1919, Czechoslovakia, Poland, and the Kingdom of the Serbs, Croats, and Slovenes.
[6]Not separately stated prior to 1891–1900.
[7]Immigrants from British North America and Mexico were not reported from 1886 to 1895, inclusive.
[8]Not separately enumerated prior to 1899.

decades from 1820 to 1920. My own Tables 9 and 10 give the distribution of the intelligence scores on the combined scale for the nativity groups we are studying. Anybody who disagrees with the estimates given in Table 33 may take these tables and split them according to any other estimates he wishes to make. However, minor changes in the proportions given in Table 33 would make very little difference in the final results. The figures which follow are merely estimates based on Table 33. I am not claiming that these figures are absolutely reliable, but merely that they represent very much closer approximations to the truth than would be obtained from the Northern and Western, and Southern and Eastern classification.

To obtain an estimate of the proportion of Nordic, Alpine, and Mediterranean blood in our immigration since 1840, the immigration figures by countries, given in Table 34, have been cut according to the proportions given in Table 33 and re-combined into percentage estimates which are given in Table 35. These estimates show in general an immigration prior to 1890 which ran 40% or 50% Nordic blood.

Since 1890, the proportion of Nordic blood has dropped to 20% or 25%, the Alpine stock now constituting about 50% of the total and the Mediterranean 20% or 25%.

The proportions given in Table 35 are shown graphically in Figure 41. The percentage estimates, given in Figure 35 and shown graphically in Figure 41, should be considered in connection with the total volume of immigration for each decade given in Table 34 and shown graphically in Figure 42.

Table No. 35

Estimate of the amount of Nordic, Alpine and Mediterranean blood coming to this country from Europe in each decade since 1840.

DECADE	TOTAL IMMIGRATION	PER CENT. NORDIC BLOOD	PER CENT. ALPINE BLOOD	PER CENT MEDITERRANEAN BLOOD	PER CENT. OTHERS AND UNCLASSIFIED
1841–1850	1,713,251	40.5	19.0	36.2	4.3
1851–1860	2,598,214	42.3	25.5	28.9	3.3
1861–1870	2,314,824	50.6	26.0	19.2	4.2
1871–1880	2,812,191	48.8	28.5	16.7	6.0
1881–1890	5,246,613	46.1	35.2	16.0	2.7
1891–1900	3,844,420	30.2	43.8	22.5	3.5
1901–1910	8,795,386	19.8	51.3	24.3	4.6
1911–1920	5,735,811	22.6	44.0	23.7	9.7

NORDIC
ALPINE
MEDITERRANEAN
OTHERS
100
90
80
70
60
50
40
30
20
10
0
1841-1850
1851-1860
1861-1870
1871-1880
1881-1890
1891-1900
1901-1910
1910-1920

Figure 41. Analysis of immigration to the United States according to the estimated amount of Nordic, Mediterranean and Alpine blood. Each bar represents 100%. The Figure was drawn from the percentages given in Table 35. Our estimates show that the Nordic immigration has fallen from about 50% to 20% or 25%, while the Alpine immigration has risen steadily from 20% or 25% to 50%. The initial decrease in the proportion of the Mediterranean immigration is due to the gradual falling off of the Irish immigration, while its subsequent recovery is due to the large numbers coming from Italy, Turkey and Greece.

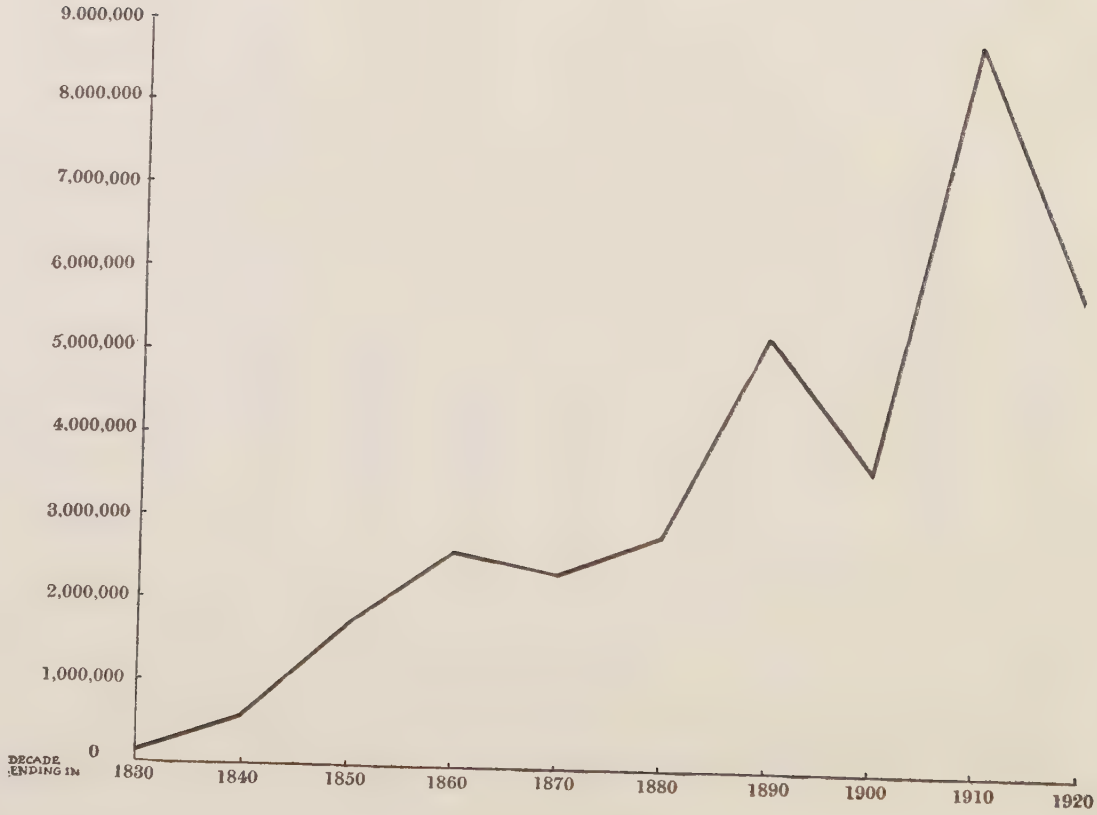
9,000,000
8,000,000
7,000,000
6,000,000
5,000,000
4,000,000
3,000,000
2,000,000
1,000,000
0
DECADE ENDING IN
1830
1840
1850
1860
1870
1880
1890
1900
1910
1920

Figure 42. Volume of immigration by decades. The percentage analysis, shown in Figure 41, should be considered in connection with the total volume of immigration shown in Tables 34 and 35 and indicated graphically above. The immigration in the decade of 1820 to 1830 was less than 150,000, in the decade 1830 to 1840, less than 600,000. Between 1901 and 1910, over 8¾ millions came to this country. In this period (1901 to 1910), Austria-Hungary, Russia and Italy sent over 5¾ millions.

In order to obtain an estimate of the intelligence of the three European races in this country, the distributions of the intelligence scores on the combined scale given in Table 9 were cut according to the proportions given in Table 33, and re-combined into Nordic, Alpine, and Mediterranean groups. The final distributions are, of course, neither purely Nordic, Alpine, nor Mediterranean, but the sample of individuals we have thus selected as Nordic is undoubtedly more typical of the Nordic race type than it is of the Alpine and Mediterranean types. In the same way, the Alpine and Mediterranean groups are more typical of each of these race types than they are of either of the other two. With thus much of apology for the method, I will, in the following pages, simply for brevity of expression, call these groups Nordic, Alpine, and Mediterranean. The reader must bear in mind that the distributions are only approximate samplings.

The actual distributions on the combined scale of the three race groups so selected are given in Table 36, together with the proportions in each thousand. The distribution curves of the three groups are shown in Figure 43, in which the horizontal direction represents scores on the combined scale, and the vertical direction proportions in each thousand making each intelligence score.

The differences found are very marked. The difference between the Nordic and Alpine group is 1.61 ± 0.042, a difference which is 38.3 times the probable error of the difference. The difference between the Nordic and Mediterranean group is 1.85 ± 0.042, a difference which is 44 times the probable error of the difference. The Alpine and Mediterranean groups are, on the other hand, very much closer together, the difference being 0.24 ± 0.04, a difference which is 6 times the probable error of the difference.

The easiest and most obvious objection that can be made

Table No. 36

Analysis of the foreign born white draft by races. Distributions of the intelligence scores of the Nordic, Alpine and Mediterranean groups.

COMBINED SCALE INTERVALS	ACTUAL DISTRIBUTION			PROPORTION IN EACH THOSUAND		
	NORDIC	ALPINE	MEDITERRANEAN	NORDIC	ALPINE	MEDITERRANEAN
24.0–24.9	....	....	....	....	....	....
23.0–23.9	....	....	....	....	....	....
22.0–22.9	3	1	1	1	....	....
21.0–21.9	8	5	2	2	1	....
20.0–20.9	19	11	5	5	2	2
19.0–19.9	37	22	11	11	5	3
18.0–18.9	71	47	26	21	10	6
17.0–17.9	135	90	55	39	19	13
16.0–16.9	238	155	103	69	32	24
15.0–15.9	357	246	180	103	51	43
14.0–14.9	469	372	296	136	78	71
13.0–13.9	566	544	468	164	114	111
12.0–12.9	528	650	591	153	136	141
11.0–11.9	371	628	590	107	132	140
10.0–10.9	260	595	569	75	125	136
9.0– 9.9	184	546	523	53	115	125
8.0– 8.9	112	403	376	32	85	90
7.0– 7.9	59	248	223	17	52	53
6.0– 6.9	26	124	108	8	26	26
5.0– 5.9	9	52	47	3	11	11
4.0– 4.9	3	19	16	1	4	4
3.0– 3.9	1	6	5	....	2	1
2.0– 2.9	....	2	1	....	....	....
1.0– 1.9	....	....	....	....	....	....
No. of cases.	3456	4766	4196			
Average....	13.28	11.67	11.43			
S.D.....	2.70	2.87	2.70			

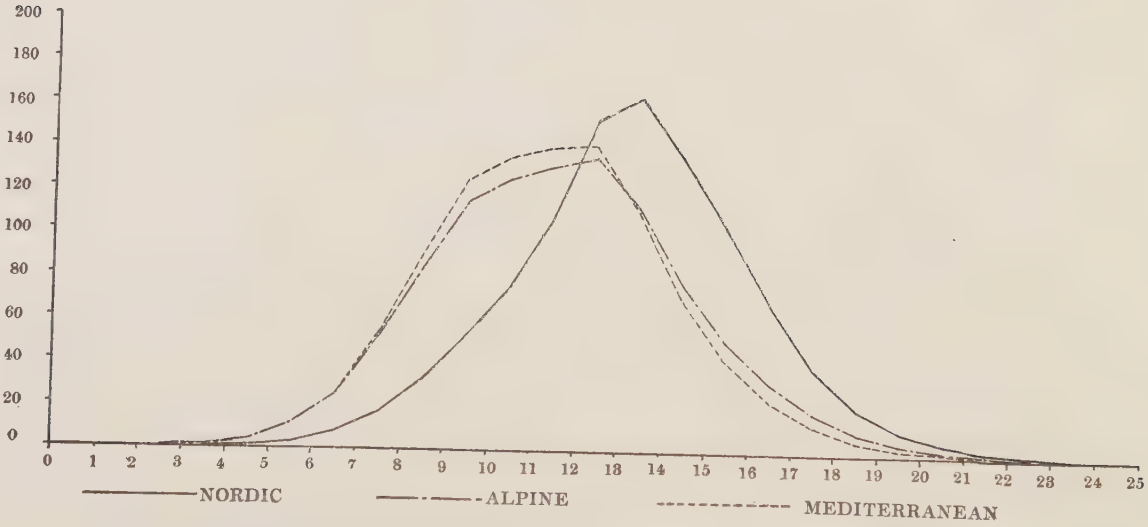

Figure 43. The distributions of the intelligence scores of the Nordic, Mediterranean and Alpine groups. This chart indicates clearly the superiority of the Nordic group. 73.9% of the Nordic group are above the average Alpine, and 76.5% of the Nordic group are above the average Mediterranean. The Mediterranean and Alpine groups are very similar, 52.3% of the Alpines exceeding the average Mediterranean.

to these findings is that the superiority of the Nordic group is due to the fact that it contains so many English speaking persons, and that lack of facility in the use of English is a handicap to the non-English speaking foreign born in the army tests. We have previously examined this hypothesis in connection with the argument establishing the fact that each succeeding five year period since 1902 shows a gradual deterioration in the intelligence of the immigrants examined in the army, and have definitely shown that the language factor does not distort the scores of the years of residence groups. There is, however, a considerable amount of wishful thinking on the subject of race, and it is well to make assurance doubly sure by testing the hypothesis that the superiority of the Nordic group is caused by the presence in the group of English speaking populations.

It is possible to split the Nordic distribution in such a way that one group will contain representatives from countries which are predominantly English speaking (England, Scotland, Ireland and Canada), while the other group will contain representatives from countries which are predominantly non-English speaking (Holland, Denmark, Germany, Sweden, Norway, Belgium, Austria, Russia, Italy and Poland). This we have done, and the results are given in Table 37, the two distributions being shown in Figure 44.

The distributions of the English speaking Nordic group and the non-English speaking Nordic group show a difference of 0.87 ± 0.065, a difference which is 13.4 times the probable error of the difference. There are, of course, cogent historical and sociological reasons accounting for the inferiority of the non-English speaking Nordic group. On the other hand, if one wishes to deny, in the teeth of the facts, the superiority of the Nordic race on the ground that the language factor mysteriously aids this group when tested,

Table No. 37

Analysis of the total Nordic sample into an English speaking Nordic group and a non-English speaking Nordic group.

COMBINED SCALE INTERVALS	ACTUAL DISTRIBUTION		PROPORTION IN EACH THOUSAND	
	ENGLISH SPEAKING NORDIC	NON-ENGLISH SPEAKING NORDIC	ENGLISH SPEAKING NORDIC	NON-ENGLISH SPEAKING NORDIC
24.0–24.9	....	....	....	....
23.0–23.9	1	....	....	....
22.0–22.9	2	....	2	....
21.0–21.9	7	2	6	1
20.0–20.9	12	6	10	3
19.0–19.9	21	16	17	7
18.0–18.9	39	32	32	14
17.0–17.9	67	67	54	30
16.0–16.9	108	131	87	59
15.0–15.9	143	214	116	96
14.0–14.9	176	293	143	132
13.0–13.9	201	365	163	164
12.0–12.9	172	356	139	160
11.0–11.9	109	262	88	118
10.0–10.9	70	189	57	85
9.0– 9.9	49	135	40	61
8.0– 8.9	31	82	25	37
7.0– 7.9	16	43	13	19
6.0– 6.9	7	19	6	9
5.0– 5.9	2	7	2	3
4.0– 4.9	1	2	....	2
3.0– 3.9	....	1	....	....
2.0– 2.9	....	....	....	....
1.0– 1.9	....	....	....	....
No. of cases...	1234	2222		
Average.....	13.84	12.97		
S. D.....	2.79	2.60		

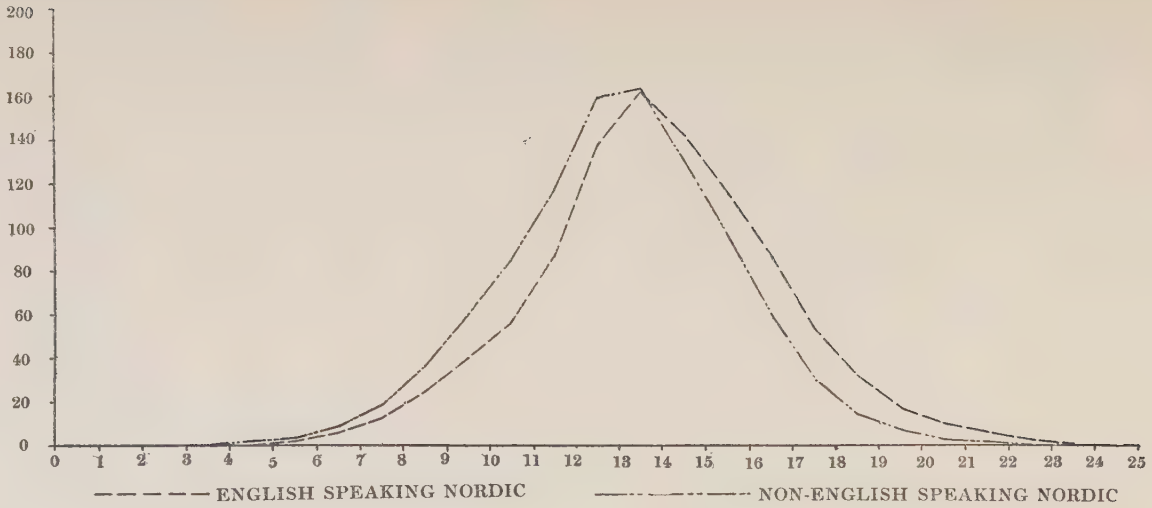

Figure 44. The distributions of the intelligence scores of the English speaking Nordic group and the non-English speaking Nordic group. These distributions were obtained by splitting the total Nordic group into two parts. The difference shown here is significant (63.4% of the English speaking Nordic group exceeding the average of the non-English speaking Nordic group), but not as large as the difference between the non-English speaking Nordic group and the Mediterranean and Alpine groups shown in Figure 45.

he may cut out of the Nordic distribution the English speaking Nordics, and still find a marked superiority of the non-English speaking Nordics over the Alpine and Mediterranean groups. The difference between the non-English speaking Nordic group and the Alpine group is 1.30 ± 0.047, a difference which is 27.6 times the probable error of the difference. The difference between the non-English speaking Nordic group and the Mediterranean group is 1.54 ± 0.047, a difference which is 31.3 times the probable error of the difference. The distributions are shown graphically in Figure 45. Discarding the English speaking Nordics entirely, we still find tremendous differences between the non-English speaking Nordic group and the Alpine and Mediterranean groups, a fact which clearly indicates that the underlying cause of the nativity differences we have shown is race, and not language.

It may be convenient for some to interpret the differences found between the representatives of the three European races in this country in terms of the standards having popular significance which were used in Section VI. The criteria of the per cent. A and B, and the per cent. D, D− and E give the following results:

	PER CENT. A AND B	PER CENT. D,D− AND E
English speaking Nordic	12.3	19.9
Total Nordic	8.1	25.8
Non-English speaking Nordic	5.7	29.1
Alpine	3.8	50.3
Mediterranean	2.5	53.6

The criteria of the per cent. at or above the average white officer, and at or below the average of the negro draft give the following results:

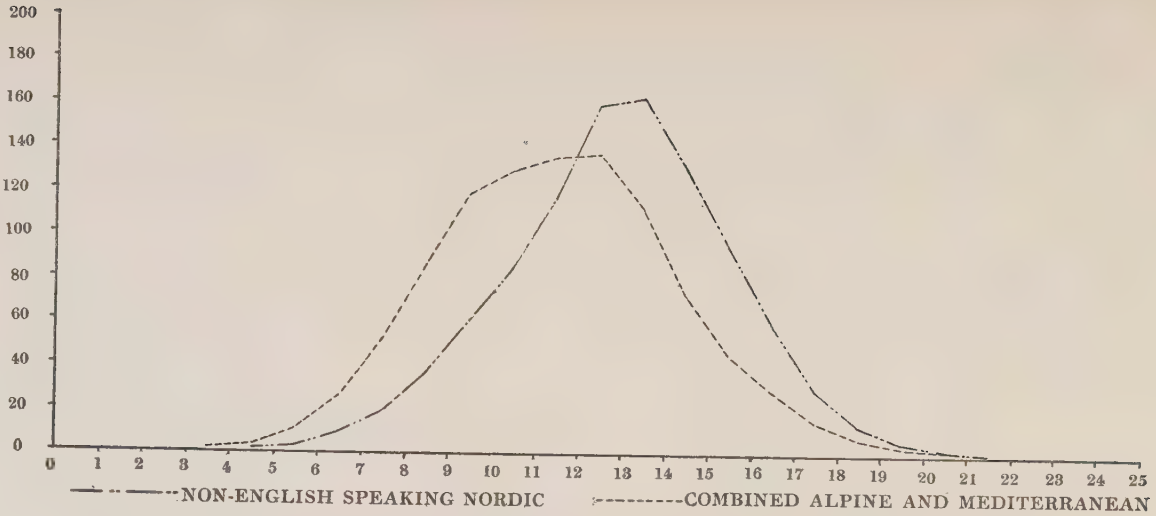

Figure 45. The distributions of the intelligence scores of the non-English speaking Nordic group and the combined Alpine and Mediterranean groups. The differences are very significant. 70.5% of the non-English speaking Nordic group exceed the average Alpine, and 73.3% exceed the average Mediterranean. If one is inclined to believe that the language factor will account for the superiority of the total Nordic group, he may disregard the English speaking Nordic group entirely, and still find the non-English speaking Nordic group superior to the Alpine and Mediterranean groups, proving that the race differences are due to intelligence and not to facility in the use of English.

	PER CENT. AT OR ABOVE AVERAGE WHITE OFFICER	PER CENT. AT OR BELOW AVERAGE OF THE NEGRO DRAFT
English speaking Nordic	4.0	10.9
Total Nordic	2.3	14.5
Non-English speaking Nordic	1.3	16.5
Alpine	1.0	34.5
Mediterranean	0.5	36.5

The criterion of the per cent. below an approximate "mental age" of eight gives the following results:

	PER CENT. BELOW "MENTAL AGE" 8
English speaking Nordic	0.8
Total Nordic	1.1
Non-English speaking Nordic	1.3
Alpine	4.2
Mediterranean	4.2

SECTION IX

RE-EXAMINATION OF PREVIOUS CONCLUSIONS IN THE LIGHT OF THE RACE HYPOTHESIS

It is now necessary to retrace our steps for a moment to examine some of our previous conclusions in the light of this new hypothesis. The hypothesis that the differences between the nativity groups found in the army tests are due to the race factor may be used to re-test our previous conclusions that each succeeding five year period of immigration since 1902 has given us an increasingly inferior selection of individuals (Section IV). The periods which we sample by means of the army data, and the average score on the combined scale of each sample are as follows:

PERIOD	NUMBER OF CASES	COMBINED SCALE AVERAGE
1887–1897	764	13.82
1898–1902	771	13.55
1903–1907	1897	12.47
1908–1912	4287	11.74
1913–1917	3576	11.41

Table 35, which gives our estimates of the per cent. of Nordic, Alpine and Mediterranean blood coming to this country, shows that the big change in immigration came between the decades 1881–1890 and 1891–1900, the percentage of Nordic blood which formerly ran from 40% to 50% having dropped to 30% in the decade 1891–1900, and to approximately 20% or 25% in the two subsequent decades. On the other hand, the big drop in the intelligence of immigrants arriving came after 1902. The change in

character of the immigration would account for part of the decline in the average intelligence of succeeding periods of immigration, but not for all of it. The decline in intelligence is due to two factors, the change in the races migrating to this country, and to the additional factor of the sending of lower and lower representatives of each race.

The only tendency which would relieve this deplorable situation would be a current of emigration strong enough to counteract the current of immigration. Table 6 preceding shows the ratio between emigration and immigration for each of the nativity groups involved in this study, and we find in general between 1908 and 1917 a return current approximately one third of the arriving current.

Unfortunately, no emigration statistics are available prior to 1908, and the figures after 1912 are distorted by the Balkan and European wars. The only sample that we can take that is comparatively free from outside influences is the sample 1908–1912. Taking the figures of arrivals and departures for this period, and dividing them into Nordic, Alpine and Mediterranean groups according to the method previously outlined, we obtain the following percentage estimates:

	ALIEN IMMIGRANTS ADMITTED	ALIEN EMIGRANTS DEPARTED	NET IMMIGRATION
Per cent. of Nordic blood	21.2	16.0	23.9
Per cent. of Alpine blood	50.4	50.6	50.2
Per cent. of Mediterranean blood	23.2	28.6	20.5
Per cent. others and unclassified	5.2	4.8	5.4

The sample from this five year period shows a slight change (approximately 3%) in favor of the Nordic type and against the Mediterranean type, the Alpine immigration holding its own. There is therefore no relief from our receding curve of intelligence from emigration, if this five year period be taken as typical of the outward alien passenger movement in other years.

It will be remembered that the army authors tentatively offered the hypothesis that the more intelligent immigrants remained in this country, while the more stupid ones went home, as a possible method of accounting for the increase of intelligence scores with increasing years of residence. The gain of 3% in favor of the Nordic immigration would produce a very slight tendency in this direction, but not enough to account for the actual increase of intelligence scores found with increasing years of residence, 11.41 (1913–1917) to 13.82 (1887–1897).

It will also be remembered that the army writers offered the hypothesis of the better adaptation of the more thoroughly Americanized group to the situation of the examination to account for the increases shown. The factor of the adaptation to the situation of the examination cannot be dissected out of the total scores of the test. If such a factor were present, it would fall equally heavily on Nordic, Alpine and Mediterranean alike, unless the change in the character of immigration were so complete that the groups sampled at the two extremes of the residence groups (1887–1897 and 1913–1917) represented different race groups.

But the difference between these two years of residence groups (2.41 ± 0.0735) is so marked that it would be necessary to assume (if our Nordic group were the more thoroughly Americanized) that the 1887–1897 group was composed entirely of English speaking Nordics or their equivalent in intelligence, and that our 1913–1917 group was

composed entirely of Mediterraneans or their equivalent in intelligence, assumptions quite unwarranted in view of the fact that in the two years of residence groups 1887–1897 and 1898-1902 we sampled 1545 individuals, while our Nordic group includes 3456 cases, and also in view of the fact that the Nordic immigration has dropped, in the period observed, at the outside from 45% to 20%. We may therefore conclude that the intangible factor of "the more thoroughly Americanized group" can not be used to explain the high test record of the Nordic group.

There is only one other possible escape from the conclusion that our test results indicate a genuine intellectual superiority of the Nordic group over the Alpine and Mediterranean groups, and that is the assumption that the situation of the examination involved a situation that was "typically Nordic." This assumption of course lands us in a perfect circle of reasoning. It would leave us with the conclusion that there was something mysteriously Nordic about alpha and beta that favored this race. We should have to assume that the Nordic, no matter where he may be, in the Canadian Northwest, in the Highlands of Scotland, or on the shores of the Baltic, is always ready for an intelligence test. Perhaps it would be easier to say that the Nordic is intelligent. A situation "typically Nordic" could not be used, however, to account for the slight but real difference between the English speaking Nordic and the non-English speaking Nordic groups. It is therefore best to abandon the attempt to account for the differences by the more or less feeble hypotheses that would make these differences an artifact of the method of examining, and recognize the fact that we are dealing with real differences in the intelligence of immigrants coming to our shores.

We have previously noted the fact that the foreign born in the army sampled as representative of the immigrants

coming to this country between 1887 and 1897 were statistically identical with the native born white draft. The change in the character of our immigration came between 1890 and 1900. The real drop in the curve of intelligence, however, started about 1900. We, therefore, cannot account for the drop in the intelligence of the immigrants sampled as representatives of those coming to this country in each five year period since 1902 by the race hypothesis entirely.

SECTION X

COMPARISON OF OUR RESULTS WITH THE CONCLUSIONS OF OTHER WRITERS ON THE SUBJECT

In a very definite way, the results which we obtain by interpreting the army data by means of the race hypothesis support Mr. Madison Grant's[1] thesis of the superiority of the Nordic type: "The Nordics are, all over the world, a race of soldiers, sailors, adventurers, and explorers, but above all, of rulers, organizers, and aristocrats in sharp contrast to the essentially peasant and democratic character of the Alpines. The Nordic race is domineering, individualistic, self-reliant, and jealous of their personal freedom both in political and religious systems, and as a result they are usually Protestants. Chivalry and knighthood and their still surviving but greatly impaired counterparts are peculiarly Nordic traits, and feudalism, class distinctions, and race pride among Europeans are traceable for the most part to the north." (p. 228.) "The pure Nordic peoples are characterized by a greater stability and steadiness than are mixed peoples such as the Irish, the ancient Gauls, and the Athenians, among all of whom the lack of these qualities was balanced by a correspondingly greater versatility." (pp. 228–229.)

Our results based on the army data also support Mr. Grant's estimates of the Alpine race: "The Alpine race is always and everywhere a race of peasants, an agricultural and never a maritime race. In fact they only extend to salt water at the head of the Adriatic and, like all purely

[1]Madison Grant. *The Passing of the Great Race.* New York, 1922, p. 476.

agricultural communities throughout Europe, tend toward democracy, although they are submissive to authority both political and religious, being usually Roman Catholics in western Europe. This race is essentially of the soil, and in towns the type is mediocre and bourgeois." (p. 227.)

Our results also support de Lapouge[1] in his contention that the Nordic type is superior to the Alpine. He says concerning the Alpine: "Il est le parfait esclave, le serf idéal, le sujet modèle, et dans les républiques comme la nôtre, le citoyen le mieux vu, car il tolère tous les abus." (p. 233.) "Les états brachycéphales, France, Autriche, Turquie, sans parler de la Pologne qui n'est plus, sont loin d'offrir la vitalité des Etats-Unis ou de l'Angleterre. Cependant la médiocrité même du brachycéphale est une force. Ce neutre échappe à toutes les causes de destruction. Noiraud, courtaud, lourdaud, le brachycéphale règne aujour d'hui de l'Atlantique à la Mer Noire. Comme la mauvaise monnaie chasse l'autre, sa race a supplanté la race meilleure. Il est inerte, il est médiocre, mais se multiplie. Sa patience est au-dessus des épreuves; il est sujet soumis, soldat passif, fonctionnaire obéissant. Il ne porte pas ombrage, il ne se révolte point." (p. 481.)

It must, however, be frankly admitted that our results, which show the Mediterranean race inferior to the Alpine, are in contradiction with those of most writers who have inferred the intellectual level of a race from its historical achievements. Mr. Grant, for instance, says: "The mental characteristics of the Mediterranean race are well known, and this race, while inferior in bodily stamina to both the Nordic and the Alpine, is probably the superior of both, certainly of the Alpines, in intellectual attainments. In the field of art its superiority to both the other European races

[1]Georges Vacher de Lapouge. *L'Aryen, son rôle social.* Paris, 1899, p. 563.

is unquestioned, although in literature and in scientific research and discovery the Nordics far excel it." (P. 229).[1]

The apparent contradiction between our results and the estimates of other observers has a very obvious solution, viz., that those who draw their conclusions from historical data are studying the Mediterranean race as it was at the period of its greatest development, when it produced the civilizations of Egypt, Mesopotamia, Phoenicia, and Crete, and, with a Nordic predominance, gave the civilizations of Greece and Rome, while our data sample this race group as it is at the present time.

The sample we have taken as representative of the Mediterranean race, as it is now constituted, is drawn from immigrants in our army born, for the most part, in Greece, Ireland, Italy, and Turkey, and inasmuch as the number from Italy (4009) is so large, our Mediterranean sample is heavily weighted (approximately 2/3) by this nativity group.

In regard to the Irish, Mr. Madison Grant says: "In spite of the fact that Paleoliths have not been found there, some indications of Paleolithic man appear in Ireland, both as single characters and as individuals. Being, like Brittany situated on the extreme western outposts of Eurasia, it has more than its share of generalized and low types surviving in the living populations, and these types, the Firbolgs, have imparted a distinct and very undesirable aspect to a large portion of the inhabitants of the west and south and have greatly lowered the intellectual status of the population as a whole. The cross between these elements and the

[1]The quotations I have chosen from Mr. Madison Grant's chapter on *Racial Aptitudes* most certainly do not do justice to that author, but they seemed to me to summarize his general position briefly. The entire book should be read to appreciate the soundness of Mr. Grant's position and the compelling force of his arguments.

Nordics appears to be a bad one, and the mental and cultural traits of the aborigines have proved to be exceedingly persistent and appear especially in the unstable temperament and the lack of coordinating and reasoning power, so often found among the Irish. To the dominance of the Mediterraneans mixed with Pre-Neolithic survivals in the south and west are to be attributed the aloofness of the island from the general trend of European civilization and its long adherence to ancient forms of religion and even to Pre-Christian superstitions." (pp. 202–203.)

The immigrants in this country from Italy come mostly from Southern Italy and Sicily. The following quotation from Ripley[1] concerning Sicily is significant:

"Commanding both straits at the waist of the Mediterranean, it has been, as Freeman in his masterly description puts it, 'the meeting place of the nations.' Tempting, therefore, and accessible, this island has been incessantly overrun by invaders from all over Europe—Sicani, Siculi, Fenicii, Greeks, and Romans, followed by Albanians, Vandals, Goths, Saracens, Normans, and last by the French and Spaniards. Is it any wonder that its people are less pure in physical type than the Sardinians or even the Calabrians of the mainland near by? Especially is this noticeable on its southern coasts, always more open to colonization than the northern edge. Nor is it surprising, as Freeman rightly adds, that 'for the very reason that Sicily has found dwelling places for so many nations, a Sicilian nation there never has been.'" (p. 271.)

The secret of the whole dilemma is the intermingling of races around the Mediterranean littoral in the last 2500 years. It is beside the point to contrast our results obtained by the actual psychological measurements of living repre-

[1]William Z. Ripley. *The Races of Europe.* New York, 1899, p. 624.

sentatives of this race with the attainments of the temporary civilizations that flared up in historical times. The whole question of the degeneration of these peoples has been discussed by Mr. Charles W. Gould,[1] and our results from the examinations of drafted men born in these regions support his position.

It is rather difficult to compare our results from the race groups with the various hypotheses erected by Professor William McDougall,[2] who, while he does not claim for the Nordic race "any general innate superiority" (p. 29), analyzes the mental constitution of this race and the other European races in such a way that an examination of his theories will be interesting. Professor McDougall's hypotheses, very briefly and inadequately stated are: that the Nordic is stronger in the instinct of curiosity, the root of wonder, than the Mediterranean; the herd instinct, the root of sociability, is stronger in the Mediterranean than in the Nordic; the Nordic is constitutionally introvert, the Mediterranean constitutionally extrovert; the instinct of self-assertion is strong in the Nordic; the Alpine is introvert but not so strongly introvert as the Nordic; the Alpine has a high degree of sociability, is perhaps relatively weak in curiosity, and strong in the instinct of submission.

In discussing innate differences in instinctive endowment, psychologists are still more or less in a speculative realm, but the field is open to experimental attack, and a body of knowledge based on experimentation is gradually growing. At the present time we must rely on concensus of opinion rather then experiment. A census of text-books on psychology would show "curiosity" usually listed as an instinctive

[1]Charles W. Gould. *America, a Family Matter*. New York, 1922, p. 196.
[2]William McDougall. *Is America Safe for Democracy*. New York, 1921, p. 213.

tendency. If we follow Professor Thorndike[1] in his analysis, and eliminate many tendencies that others include, we shall still have left the instincts of multiform mental and physical activity as the potent movers of men's economic and recreative activities. (p. 144.) If any instinctive tendency finds expression in the tasks assigned by the army tests, it is this instinct for multiform mental activity, more vaguely termed "curiosity." Our tests, however, measure the end result of such a tendency and not the tendency itself, and it is only in this vague way that our results showing the definite intellectual superiority of the Nordic race can be taken as substantiating or contributing to Professor McDougall's hypothesis.

It is difficult to check our results from the analysis of the foreign born white draft by country of birth (reported in Section VI) with the results of other investigators, on account of the different tests that were used, and the different methods of selecting subjects. Miss Murdoch[2] examined, by means of the Pressey group point scale, 500 Jewish children and 500 Italian children at one school in New York City, and 500 American children and 230 negro children at another school. The American and Jewish children tested about the same. About 15% of the Italians equalled or exceeded the median of the Jews, and about 30% of the negroes equalled or exceeded the median of the Jews. The investigation equalizes the environmental factor by selecting, in one instance, Italians and Jews from the same school and consequently from the same general neighborhood (East 110th St. near 2nd Ave.), and, in the other instance, by selecting native white and negro chil-

[1]E. L. Thorndike, *Educational Psychology*, Vol. I. *The Original Nature of Man*, New York, 1919, p. 327.
[2]K. Murdoch. *A Study of Race Differences in New York City*. School and Society, 1920, pp. 11, 147–150.

dren from the same general neighborhood (West side, 8th Ave., near 140th St.), but the American children living in this neighborhood can not be taken as typical of the country as a whole.

Miss Arlitt[1] concludes from her examination of 343 children, (191 native born Americans, 87 Italians, and 71 negroes), by the Stanford-Binet scale, that "there is a marked difference in the distribution of intelligence in groups of the same race but different social status," and states that "race norms which do not take the social status factor into account are apt to be to that extent invalid." (p. 183.) This position seems to ignore the observation, repeatedly confirmed by experiment, that children from the professional, semi-professional and higher business classes have, on the whole, an hereditary endowment superior to that of children from the semi-skilled and unskilled laboring classes. Terman[2] states "It has in fact been found wherever comparisons have been made that children of superior social status yield a higher average mental age than children of the laboring classes. . . . However, the common opinion that the child from a cultured home does better in tests solely by reason of his superior home advantages is an entirely gratuitous assumption. Practically all of the investigations which have been made of the influence of nature and nurture on mental performance agree in attributing far more to original endowment than to environment. Common observation would itself suggest that the social class to which the family belongs depends less on chance than on the parents' native qualities of intellect and character." (p. 115.)

[1]A. H. Arlitt. *On the Need for Caution in Establishing Race Norms.* Journal of Applied Psychology, 1921, pp. 5, 179–183.
[2]L. M. Terman. *The Measurement of Intelligence.* Boston, 1916, p. 362.

One frequently hears the opinion expressed in scientific circles that differences found between racial groups can not be attributed to race unless the individuals examined are drawn from the same social *milieu*. Miss Arlitt finds native born white children of inferior and very inferior social status above the Italian and negro children in intelligence, but attributes the larger differences found between the entire native white group and the Italian and negro groups to the fact that three eighths of the native white children come from homes of superior and very superior social status. In the same way, Miss Murdoch finds Jews living near East 110th St. and 2nd Ave. in New York City not very inferior to native born whites living in the mixed white and negro section around 8th Ave. and 140th St. The equalization of the environmental factor is a necessary control in certain phases of scientific experiments on race differences, but conclusions as to the intelligence of racial groups must be drawn from samples taken at random from the entire country. These conditions are more nearly met by the army sampling of individuals in the draft. Our samples of 81,465 native born individuals in the white draft, of 12,492 foreign born individuals, and 23,596 negroes are drawn impartially from every section of the country. If we selected our native born Americans from those who live in the same squalid conditions in which we find most of our negro and foreign born population, we would not have a fair sample.

It is unfortunate that our army data classify foreign born individuals only by country of origin, so that we have no separate intelligence distributions for the Jews. According to the 1910 census, about 50% of the foreign born population reporting Russia as their country of origin spoke Hebrew or Yiddish, about 25% spoke Polish, less than 3% spoke Russian, and the rest spoke Lithuanian, Lettish,

German, Finnish, Ruthenian and other tongues. From the immigration statistics showing aliens admitted classified according to race or people, we find about 10% (arriving between 1900 and 1920) reported as Hebrew. It is fair to assume that our army sample of immigrants from Russia is at least one half Jewish, and that the sample we have selected as Alpine[1] is from one fifth to one fourth Jewish.

Our figures, then, would rather tend to disprove the popular belief that the Jew is highly intelligent. Immigrants examined in the army, who report their birthplace as Russia, had an average intelligence below those from all other countries except Poland and Italy. It is perhaps significant to note, however, that the sample from Russia has a higher standard deviation (2.83) than that of any other immigrant group sampled, and that the Alpine group has a higher standard deviation than the Nordic or Mediterranean groups (2.60). If we assume that the Jewish immigrants have a low average intelligence, but a higher variability than other nativity groups, this would reconcile our figures with popular belief, and, at the same time, with the fact that investigators searching for talent in New York City and California schools find a frequent occurence of talent among Jewish children, The able Jew is popularly recognized not only because of his ability, but because he is able and a Jew.

Our results showing the marked intellectual inferiority of the negro are corrobated by practically all of the investigators who have used psychological tests on white and negro groups. This inferiority holds even when a low intellectual sampling of whites is made by selecting only

[1]There is no serious objection, from the anthropological standpoint, to classifying the northern Jew as an Alpine, for he has the head form, stature, and color of his Slavic neighbors. He is an Alpine Slav.

those who live in the same environment, and who have had the same educational opportunities. Professor Ferguson,[1] who has studied the problem most carefully, concludes that in general 25% of the negroes exceed the median white. Our figures show a greater difference than he estimates, less than 12% of the negroes exceeding the average of the native born white draft. Professor Ferguson also estimates that 20% of pure negroes, 25% of negroes three quarters pure, 30% of the true mulattoes, and 35% of the quadroons equal or exceed the average score of *comparable* whites.

The discrepancies between data from various investigators as to the *amount* of difference between negroes and whites probably result from different methods of selecting whites. If we compare negroes only to those whites who live in the same neighborhood, and who have had the same educational opportunities, our differences are smaller than those obtained by comparing samples of the entire white and negro populations.

Some writers would account for the differences found between white and negro by differences of educational opportunity alone. The army tests showed the northern negro superior to the southern negro, and this superiority is attributed to the superior educational opportunities in the North. The educational record of the negro sample we are studying shows that more than half of the negroes from the southern States did not go beyond the third grade, and only 7% finished the eighth grade, while about half of the northern negroes finished the fifth grade, and a quarter finished the eighth grade. That the difference between the northern and southern negro is not entirely due to school-

[1]G. O. Ferguson. *The Mental Status of the American Negro.* Scientific Monthly, 1921, pp. 12, 533–543.

ing, but partly to intelligence, is shown by the fact that groups of southern and northern negroes of *equal schooling* show striking differences in intelligence.

The superior intelligence measurements of the northern negro are due to three factors: first, the greater amount of educational opportunity, which does affect, to some extent, scores on our present intelligence tests; second, the greater amount of admixture of white blood; and, third, the operation of economic and social forces, such as higher wages, better living conditions, identical school privileges, and a less complete social ostracism, tending to draw the more intelligent negro to the North. It is impossible to dissect out of this complex of forces the relative weight of each factor. No psychologist would maintain that the mental tests he is now using do not measure educational opportunity to some extent. On the other hand, it is absurd to attribute all differences found between northern and southern negroes to superior educational opportunities in the North, for differences are found between groups of the same schooling, and differences are shown by beta as well as by alpha.

At the present stage of development of psychological tests, we can not measure the actual *amount* of difference in intelligence due to race or nativity. We can only prove that differences do exist, and we can interpret these differences in terms that have great social and economic significance. The intellectual superiority of our Nordic group over the Alpine, Mediterranean, and negro groups has been demonstrated. If a person is unwilling to accept the race hypothesis as developed here, he may go back to the original nativity groups, and he can not deny the fact that differences exist.

When our methods of measuring intellectual capacity have been perfected, we will be in a position to determine

quantitatively the amount of race differences. Rough group tests of the type we are now using will indicate the fact that differences exist. However, while scientists are perfecting their methods of examining, it would be well for them to perfect their logic at the same time. Particularly misleading and unsound is the theory that disregards all differences found between racial groups unless the groups have had the same educational and environmental opportunities.

This theory in its most extreme form is set forth by Garth[1] as follows:

"The elements in a study of racial mental similarities or differences must be these: (1) Two so-called races R_1 and R_2, (2) an equal amount of educational opportunity, E, which should include social pressure and racial patterns of thought, and (3) psychological tests, D, within the grasp of both racial groups. We should have as a result of our experiment R_1 E D equal to, greater than, or less than R_2 E D. In this experiment the only unknown elements should be R_1 and R_2. If E could be made equal the experiment could be worked.

"This element of educational opportunity-nurture, is the one causing most of the trouble in racial psychology as an uncontrollable element. It does not offer quite so much difficulty in the study of sex differences, yet it is there only in smaller degree than in racial differences, and as it is controlled the 'sex differences' tend to disappear. Since this element of education, or nurture, cannot be eliminated it would be safer to take for comparison such racial groups as have had as nearly the same educational opportunity as is possible having any disparity of this sort well in mind when we interpret the results of the experiment. Having

[1]T. R. Garth. *White, Indian and Negro Work Curves.* Journal of Applied Psychology, 1921, *5*, 14–25.

done this, we first take the complete distributions on the scale of measurement for the groups as statements of the true facts of the case, race for race. We then combine these distributions into a total distribution of accomplishment of all the races taken together to see if we have multimodal effects. Should we find these effects we may conclude that we have evidence of types, or racial types, and there should in this case be one mode for each racial group. But should the combined distribution for the several racial groups reveal only one mode we may conclude that the test reveals no types—no real racial differences but rather similarities." (p. 16.)

If intelligence counts for anything in the competition among human beings, it is natural to expect that individuals of superior intelligence will adjust themselves more easily to their physical and social environment, and that they will endow their children not only with material goods, but with the ability to adjust themselves to the same or a more complex environment. To select individuals who have fallen behind in the struggle to adjust themselves to the civilization their race has built as typical of that race is an error, for their position itself shows that they are, for the most part, individuals with an inferior hereditary endowment.

In the same way, our educational institutions are themselves a part of our own race heritage. The average negro child can not advance through an educational curriculum adapted to the Anglo-Saxon child in step with that child. To select children of equal education, age for age, in the two groups, is to sample either superior negroes or inferior whites.

The scientific problem is that of eliminating from the tests used as measuring instruments those particular tests which demonstrably measure nurture, and to measure,

with genuine tests of native intelligence, random or impartial samples from each race throughout the entire range of its geographical and institutional distribution.

CONCLUSIONS

Our study of the army tests of foreign born individuals has pointed at every step to the conclusion that the average intelligence of our immigrants is declining. This deterioration in the intellectual level of immigrants has been found to be due to two causes. The migrations of the Alpine and Mediterranean races have increased to such an extent in the last thirty or forty years that this blood now constitutes 70% or 75% of the total immigration. The representatives of the Alpine and Mediterranean races in our immigration are intellectually inferior to the representatives of the Nordic race which formerly made up about 50% of our immigration. In addition, we find that we are getting progressively lower and lower types from each nativity group or race.

In the light of our findings in Sections IV and IX, it is possible to re-draw our curve (Figure 33) representing increase of intelligence score with increasing years of residence and to represent it truly as in Figure 46, which shows the decline of intelligence with each succeeding period of immigration.

It is also possible to make a picture of the elements now entering into American intelligence. At one extreme we have the distribution of the Nordic race group. At the other extreme we have the American negro. Between the Nordic and the negro, but closer to the negro than to the Nordic, we find the Alpine and Mediterranean types. These distributions we have projected together in Figure 47.

Throughout this study all measurements have been made in terms of averages and variability about the average. In interpreting averages, we must never forget that they stand

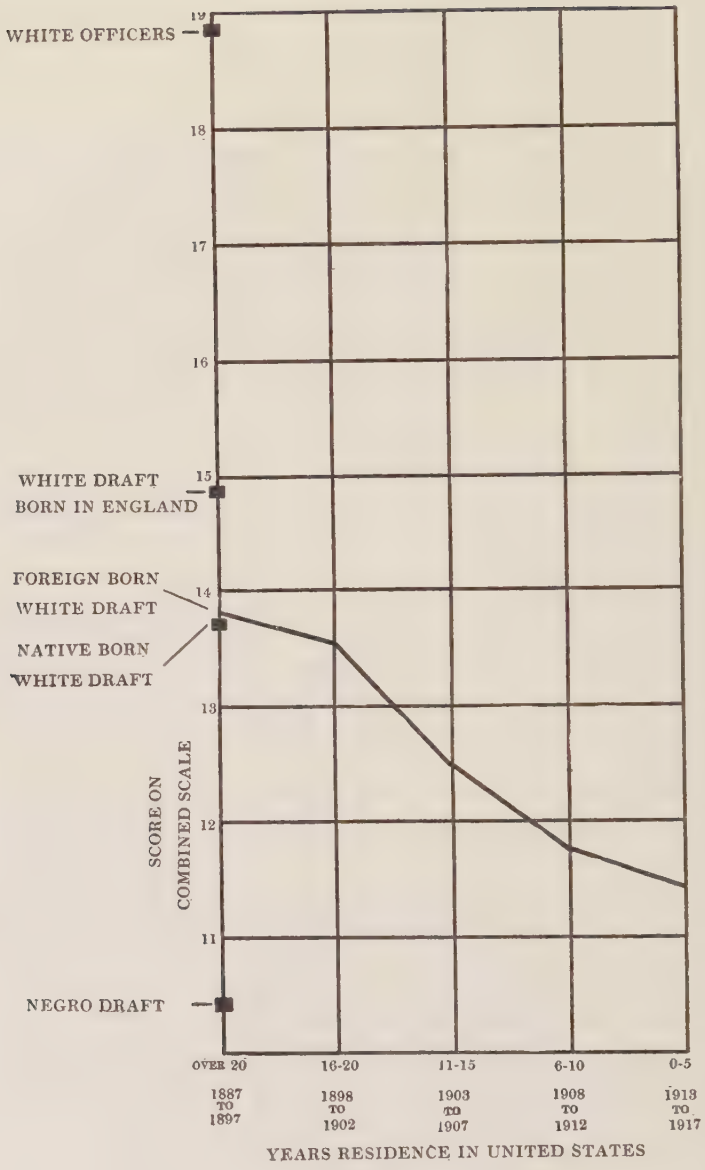
WHITE OFFICERS
19
18
17
16
WHITE DRAFT
BORN IN ENGLAND
15
FOREIGN BORN
WHITE DRAFT
14
NATIVE BORN
WHITE DRAFT
13
12
11
NEGRO DRAFT
SCORE ON
COMBINED SCALE
OVER 20
16-20
11-15
6-10
0-5
1887
TO
1897
1898
TO
1902
1903
TO
1907
1908
TO
1912
1913
TO
1917
YEARS RESIDENCE IN UNITED STATES

Figure 46. The decline of intelligence with each succeeding period of immigration. The apparent increase of intelligence with increasing length of residence, as shown in Figure 33, has been proved to be a progressive decrease in the intellectual level of immigrants coming to this country in each succeeding five-year period since 1902. The evidence indicates that the immigrants prior to 1902 were intellectually equal to the native born white draft. The army sample of "native born" includes, besides native born of native parentage, the native born of foreign or mixed parentage. It is perhaps possible that the native born of native parentage might have tested higher than 13.77. The position of the white draft born in England is shown above. Although the true position of the native born American may be a matter of speculation, there is no doubt that the more recent immigrants are intellectually closer to the negro than to the native born white sample.

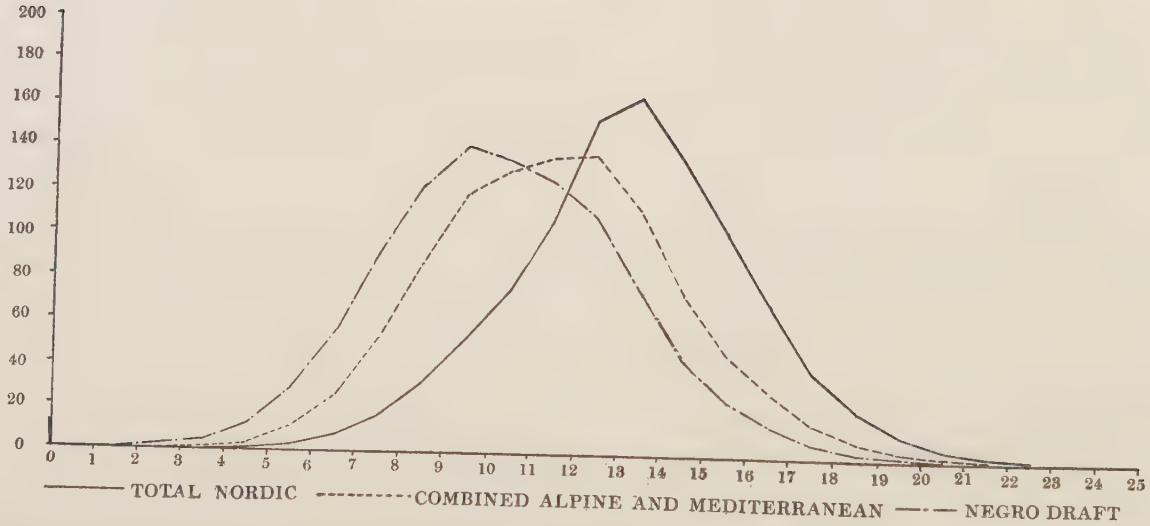
200
180
160
140
120
100
80
60
40
20
0
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25
TOTAL NORDIC
COMBINED ALPINE AND MEDITERRANEAN
NEGRO DRAFT

Figure 47. The constituent elements of American intelligence. The distributions of the intelligence scores of the entire Nordic group, the combined Mediterranean and Alpine groups, and the negro draft. The process of racial intermixture cannot result in anything but an average of these elements, with the resulting deterioration of American intelligence. About 85% of the Nordic group exceed the average negro. About 75% of the Nordic group exceed the Alpine and Mediterranean. About 35% of the Alpine and Mediterranean groups are below the average negro.

for an entire distribution. Careless thinkers are prone to select one or two striking examples of ability from a particular group, and then rest confidently in the belief that they have overthrown an argument based on the total distribution of ability. The Fourth of July orator can convincingly raise the popular belief in the intellectual level of Poland by shouting the name of Kosciusko from a high platform, but he can not alter the distribution of the intelligence of the Polish immigrant. All countries send men of exceptional ability to America, but the point is that some send fewer than others.

Our distribution curve of intelligence includes ability as well as defect. The English speaking Nordic group, for instance, averages 13.84, and furnishes at one extreme about 40 men in 1000 who are above the average white officer, while at the other extreme, the group furnishes about 8 in 1000 who are below an estimated "mental age" of eight. A distribution further down the scale contributes more to the lower orders of intelligence. The distribution of the intelligence scores of the negro draft, for instance, indicates that they contribute only 4 in 1000 above the average white officer, while they give us 100 in 1000 below the approximate "mental age" of eight. The Alpine and Mediterranean races give us only 5 or 10 in 1000 above the average ability of the white officer, and about 40 in 1000 below the "mental age" of eight. About 350 in 1000 of the Alpine and Mediterranean types are below the average negro.

The intellectual characteristics of the immigration to the United States as measured by the samples in the draft have been reported in this study, first by country of birth, and second by race. Parallel with the measurements of intelligence, the figures on immigration have been presented. To complete the picture, there is presented in Table 38 the population of the United States according to the 1920

Table No. 38

Population of the United States in 1920

Native White of Native Parentage	58,421,957	
Native White of Foreign Parentage	15,694,539	
Native White of Mixed Parentage	6,991,665	
Total Native White		81,108,161
Foreign Born White		13,712,754
Negro		10,463,131
Indian		244,437
Chinese		61,639
Japanese		111,010
All others:		
Filipinos	5,603	
Hindus	2,507	
Koreans	1,224	
Hawaiians	110	
Malays	19	
Siamese	17	
Samoans	6	
Maoris	2	
Total all others		9,488
Total Population		105,710,620

NOTE: Clinton Stoddard Burr in *America's Race Heritage* (New York: The National Historical Society, 1922, pp. 327) estimates that in 1920 there were 44,689,278 descendents of the old Colonial white stock.

census.[1] We have with us approximately 15 7/10 millions of individuals of foreign parentage, 7 millions of mixed parentage, 13 7/10 millions of foreign born, and 10 ½ millions of negroes. Roughly, in every 100 of our population, 55 are native born of native parentage, and the other 45 foreign born, or of foreign, mixed, or colored parentage. The group of native born of native parentage includes many children of the immigrants coming to this country prior to 1890.

Our immigration figures show a very decided shift from the Nordic in favor of the Alpine. The immigration between 1820 and 1890 probably never contained more than 50% or 60% Nordic blood, and prior to 1820 there was very little immigration. The earliest settlers were almost pure Nordic types, and we may assume the existence by 1820 of a race as predominantly Nordic as that of England. This recent change was, of course, reflected in the cross section of the foreign born population taken at 1910, and which constitutes the basis of our present immigration act restricting immigration to 3% of the nationals then resident here. A rough estimate of the racial composition of the quotas from various countries admissable under the new law shows about 35% Nordic blood, 45% Alpine blood and 20% Mediterranean blood in the annual stream of approximately 1/3 of a million that may enter.

There can be no doubt that recent history has shown a movement of inferior peoples or inferior representatives of

[1]Too much reliance can not be placed on the census returns for the foreign born white population. The 1910 census shows the foreign born white population as 13,345,545, while the 1920 census shows that population as 13,712,754, which gives a net increase of 367,209. On the other hand, the figures of the Commissioner General of Immigration show by actual count at the ports 5,725,811 aliens admitted and 2,146,994 aliens departed, leaving a net increase of 3,578,817 for the same period covered by the two censuses (1910 and 1920). Inasmuch as the enumerators could not have missed three million, they are probably counted among the native white.

peoples to this country. Few people realize the magnitude of this movement or the speed with which it has taken place. Since 1901, less than a single generation, it may be estimated that about 10,000,000 Alpine and Mediterranean types have come to this country. Allowing for the return of 1/3 or 3/8 of these, and using our army estimates of intellectual ability, this would give us over 2,000,000 immigrants below the average negro.

We may consider that the population of the United States is made up of four racial elements, the Nordic, Alpine, and Mediterranean races of Europe, and the negro. If these four types blend in the future into one general American type, then it is a foregone conclusion that this future blended American will be less intelligent than the present native born American, for the general results of the admixture of higher and lower orders of intelligence must inevitably be a mean between the two.

If we turn to the history of races, we find that as a general rule where two races have been in contact they have intermingled, and a cross between the two has resulted. Europe shows many examples of areas where the anthropological characteristics of one race shade over into those of another race where the two have intermixed, and, indeed, in countries such as France and Switzerland it is only in areas that are geographically or economically isolated that one finds types that are relatively pure. The Mongol-Tatar element in Russia is an integral part of the population. The Mediterranean race throughout the area of its contact with the negro has crossed with him. Some of the Berbers in Northern Africa show negroid characteristics, and in India the Mediterranean race has crossed with the Dravidians and Pre-Dravidian negroids. The population of Sardinia shows a number of negroid characteristics. Turn where we may, history gives us no great exception to the general rule

that propinquity leads to opportunity and opportunity to intermixture.

In considering racial crosses, Professor Conklin[1] states that "It is highly probable that while some of these hybrids may show all the bad qualities of both parents, others may show the good qualities of both and indeed in this respect resemble the children in any pure-bred family. But it is practically certain that the general or average results of the crossing of a superior and an inferior race are to strike a balance somewhere between the two. This is no contradiction of the principles of Mendelian inheritance but rather the application of these principles to a general population. The general effect of the hybridization of races can not fail to lead to a lowering of the qualities of the higher race and a raising of the qualities of the lower one." (pp. 50–51.)

And as to the possibility of a cross between races in the future, Professor Conklin writes: "Even if we are horrified by the thought, we cannot hide the fact that all present signs point to an intimate commingling of all existing human types within the next five or ten thousand years at most. Unless we can re-establish geographical isolation of races, we cannot prevent their interbreeding. By rigid laws excluding immigrants of other races, such as they have in New Zealand and Australia, it may be possible for a time to maintain the purity of the white race in certain countries, but with constantly increasing intercommunications between all lands and peoples such artificial barriers will probably prove as ineffectual in the long run as the Great Wall of China. The races of the world are not drawing apart but together, and it needs only the vision that will look

[1]Edwin G. Conklin. *The Direction of Human Evolution.* New York, 1921, pp. 247.

ahead a few thousand years to see the blending of all racial currents into a common stream." (p. 52.)

If we frankly recognize the fact that the crossing of races in juxtaposition has always occurred in the past, what evidence have we that such crosses have had untoward consequences? Our own data from the army tests indicate clearly the intellectual superiority of the Nordic race group. This superiority is confirmed by observation of this race in history. The Alpine race, according to our figures, which are supported by historical evidence, seems to be considerably below the Nordic type intellectually. However, our recruits from Germany, which represents a Nordic-Alpine cross, are about the same as those from Holland, Scotland, the United States, Denmark, and Canada, countries which have on the whole a greater proportion of Nordic blood than Germany. Again, the Nordic and Alpine mixture in Switzerland has given a stable people, who have evolved, in spite of linguistic differences, a very advanced form of government. The evidence indicates that the Nordic-Alpine cross, which occurred in Western Europe when the Nordics overwhelmed the Alpines to such an extent that the type was completely submerged and not re-discovered until recently, has not given unfortunate results.

This evidence, however, can not be carried over to indicate that a cross between the Nordic and the Alpine Slav would be desirable. The Alpines that our data sample come for the most part from an area peopled largely by a branch of the Alpine race which appeared late and radiated from the Carpathian Mountains. It is probably a different branch of the Alpine race from that which forms the primitive substratum of the present population of Western Europe. Our data on the Alpine Slav show that he is intellectually inferior to the Nordic, and every indication

would point to a lowering of the average intelligence of the Nordic if crossed with the Alpine Slav. There can be no objection to the intermixture of races of equal ability, provided the mingling proceeds equally from all sections of the distribution of ability. Our data, however, indicate that the Alpine Slav we have imported and to whom we give preference in our present immigration law is intellectually inferior to the Nordic type.

The Mediterranean race at its northern extension blends with the Alpine very considerably, and to a less extent with the Nordic. At the point of its furthermost western expansion in Europe it has crossed with the primitive types in Ireland. Throughout the area of its southern and eastern expansion it has crossed with negroid types. In this continent, the Mediterranean has crossed with the Amerind and the imported negro very extensively. In general, the Mediterranean race has crossed with primitive race types more completely and promiscuously than either the Alpine or the Nordic, and with most unfortunate results.

We must now frankly admit the undesirable results which would ensue from a cross between the Nordic in this country with the Alpine Slav, with the degenerated hybrid Mediterranean, or with the negro, or from the promiscuous intermingling of all four types. Granted the undesirable results of such an intermingling, is there any evidence showing that such a process is going on? Unfortunately the evidence is undeniable. The 1920 census shows that we have 7,000,000 native born whites of mixed parentage, a fact which indicates clearly the number of crosses between the native born stock and the European importations.

The evidence in regard to the white and negro cross is also indisputable. If we examine the figures showing the proportion of mulattoes to a thousand blacks for each twenty year period from 1850 to 1910, we find that in 1850

there were 126 mulattoes to a thousand blacks, 136 in 1870, 179 in 1890 and 264 in 1910. This intermixture of white and negro has been a natural result of the emancipation of the negro and the breaking down of social barriers against him, mostly in the North and West. In 1850, the free colored population showed 581 mulattoes to a thousand blacks as against 83 in the slave population. At each of the four censuses (1850, 1870, 1890 and 1910) the South, where the social barriers are more rigid than elsewhere, has returned the smallest proportion of mulattoes to a thousand blacks. The 1910 census showed 201 in the South, 266 in the North and 321 in the West, and the West has returned the highest proportion at each of the censuses except 1850.

We must face a possibility of racial admixture here that is infinitely worse than that faced by any European country to-day, for we are incorporating the negro into our racial stock, while all of Europe is comparatively free from this taint. It is true that the rate of increase of the negro in this country by ten year periods since 1800 has decreased rather steadily from about 30% to about 11%, but this declining rate has given a gross population increase from approximately 1,000,000 to approximately 10,000,000. It is also true that the negro now constitutes only about 10% of the total population, where he formerly constituted 18% or 19% (1790 to 1830), but part of this decrease in percentage of the total population is due to the great influx of immigrants, and we favor in our immigration law those countries 35% of whose representatives here are below the average negro. The declining rate of increase in the negro population from 1800 to 1910 would indicate a correspondingly lower rate to be expected in the future. From 1900 to 1920 the negro population increased 18.4%, while the native born white of native parents increased 42.6%, and the native born white of foreign parents increased 47.6%. It is

impossible to predict at the present time that the rate of infiltration of white blood into the negro will be checked by the declining rate of increase in the negro blood itself. The essential point is that there are 10,000,000 negroes here now and that the proportion of mulattoes to a thousand blacks has increased with alarming rapidity since 1850.

According to all evidence available, then, American intelligence is declining, and will proceed with an accelerating rate as the racial admixture becomes more and more extensive. The decline of American intelligence will be more rapid than the decline of the intelligence of European national groups, owing to the presence here of the negro. These are the plain, if somewhat ugly, facts that our study shows. The deterioration of American intelligence is not inevitable, however, if public action can be aroused to prevent it. There is no reason why legal steps should not be taken which would insure a continuously progressive upward evolution.

The steps that should be taken to preserve or increase our present intellectual capacity must of course be dictated by science and not by political expediency. Immigration should not only be restrictive but highly selective. And the revision of the immigration and naturalization laws will only afford a slight relief from our present difficulty. The really important steps are those looking toward the prevention of the continued propagation of defective strains in the present population. If all immigration were stopped now, the decline of American intelligence would still be inevitable. This is the problem which must be met, and our manner of meeting it will determine the future course of our national life.

COUNTWAY LIBRARY
HC 1GEA R